

TESIS DE DOCTORADO

Producción biotecnológica de proteasas aspárticas vegetales para la elaboración de nuevos productos lácteos

Lucía Feijoo Siota

ESCUELA DE DOCTORADO INTERNACIONAL EN CIENCIAS DE LA SALUD

PROGRAMA DE DOCTORADO EN AVANCES EN BIOLOGÍA MICROBIANA Y PARASITARIA

SANTIAGO DE COMPOSTELA

AÑO 2018



DECLARACIÓN DEL AUTOR DE LA TESIS

**Producción biotecnológica de proteasas aspárticas vegetales para la
elaboración de nuevos productos lácteos**

Dña. Lucía Feijoo Siota

Presento mi tesis, siguiendo el procedimiento adecuado al Reglamento, y declaro que:

- 1) *La tesis abarca los resultados de la elaboración de mi trabajo.*
- 2) *En su caso, en la tesis se hace referencia a las colaboraciones que tuvo este trabajo.*
- 3) *La tesis es la versión definitiva presentada para su defensa y coincide con la versión enviada en formato electrónico.*
- 4) *Confirmo que la tesis no incurre en ningún tipo de plagio de otros autores ni de trabajos presentados por mí para la obtención de otros títulos.*

En Santiago de Compostela, 23 de Julio de 2018

Fdo. Lucía Feijoo Siota

Conflito de intereses

Eu, Lucía Feijoo Siota, con DNI 52492952Y, declaro que son titular da seguinte solicitude de patente, desenvolvida no contexto da tese de doutoramento “Producción biotecnolóxica de proteasas aspárticas vegetales para la elaboración de produtos lácteos”.

Título: Cepa recombinante, método de produción de proteasas aspárticas de *Galium verum* y uso en la industria láctea.


Autores: González Villa, Tomás
Feijoo Siota, Lucía
Rodríguez Rama, José Luís
de Miguel Bouzas, Trinidad

Fecha de publicación: 28/06/2018

Nº de publicación: WO/2018/115559

Nº de la solicitud internacional: PCT/ES2017/070832

Santiago de Compostela, 24 de setembro do 2018



Lucía Feijoo Siota

AUTORIZACIÓN DE LOS DIRECTORES DE LA TESIS

Producción biotecnológica de proteasas aspárticas vegetales para la
elaboración de nuevos productos lácteos

D. Tomás González Villa

Dña. María Trinidad de Miguel Bouzas

INFORMAN:

*Que la presente tesis, corresponde con el trabajo realizado por Dña. **Lucía Feijoo Siota**, bajo nuestra dirección, y autorizamos su presentación, considerando que reúne los requisitos exigidos en el Reglamento de Estudios de Doctorado de la USC, y que como directores de ésta no incurre en las causas de abstención establecidas en Ley 40/2015.*

En Santiago de Compostela, 23 de Julio de 2018

Fdo. Tomás González Villa

Fdo. Trinidad de Miguel Bouzas

AGRADECIMIENTOS

Me gustaría expresar mi agradecimiento a aquellas personas que de algún modo han contribuido a la realización de este trabajo:

A mis directores de tesis, el profesor Tomás Gonzalez Villa, por permitirme formar parte de su grupo de investigación y a la profesora Trinidad de Miguel Bouzas. Gracias por todo lo que he aprendido gracias a vosotros, por vuestra disponibilidad y apoyo. Por todo lo que me habéis aportado a nivel profesional y también personal.

A Tomás, por todas esas ideas para nuevos proyectos, por la pasión que pone en su trabajo, en los papers, en los capítulos...

A todos los compañeros que han pasado por el grupo de investigación durante el desarrollo de esta tesis, en especial a Lucía Blasco y a José Luis RR, con quien he compartido un montón de experiencias; gracias por vuestra ayuda, optimismo, amistad...

A toda la gente relacionada con el mundo de la investigación que he conocido durante todos estos años. A Montse, de la AVTE, que ha puesto todo de su parte para que el resultado de este trabajo llegase al mercado (a lo mejor algún día...). Al personal de la facultad de Farmacia por su amabilidad.

A mis amigos (y a Carlos, que está de camino), por estar siempre ahí, por los ánimos, voy a echar de menos la pregunta del millón.

A mi familia, a mi abuela, a mis hermanos, Sergio e Inés y a mis sobrinas, Lulu y María, sois estupendos.

A Fran, por ser capaz de escuchar el mismo rollo mil veces y hacer como si fuese la primera, por su apoyo y todo lo que compartimos.

En especial quiero darle las gracias a mis padres, Dori y Pepe, que lo han hecho todo por mí, incluso agradecerme sus sacrificios. Sin su ayuda no podría haber realizado este trabajo. Os quiero y os admiro.



A mis padres

ÍNDICE

RESUMEN	i
SUMMARY	iii
ABREVIATURAS	v
ÍNDICE DE TABLAS	xi
ÍNDICE DE FIGURAS	xiii
1. INTRODUCCIÓN	1
1.1 Proteasas	1
1.2 Proteasas aspárticas (PAs) de plantas	3
1.2.1 Clasificación	3
1.2.2 Características comunes de las PAs	5
1.2.3 Estructura molecular de la PAs de plantas típicas	7
1.2.3.1 Estructura primaria	7
1.2.3.2 Estructura terciaria	12
1.2.4 Procesamiento y mecanismos de activación de PAs de plantas típicas.	16
1.2.5 Distribución, localización y funciones biológicas.	20
1.2.6 Aplicaciones industriales.....	25
1.3 El queso	26
1.3.1 La leche	27
1.3.2 La coagulación de la leche	28
1.3.3 Coagulantes de la leche	29
1.3.3.1 Definición	29
1.3.3.2 Tipos de cuajos y coagulantes.....	30
1.3.3.3 Consideraciones en la elección de un coagulante	35

1.3.3.4 Participación de los coagulantes en la formación del sabor	36
1.4 Investigaciones sobre coagulantes vegetales	38
1.4.1 Extractos vegetales y sus quesos	38
1.4.2 Proteasas aspárticas vegetales recombinantes y sus quesos	51
1.5 Especies vegetales objeto de estudio	53
1.5.1 <i>Galium verum</i>	53
1.5.2 <i>Ficus carica</i>	55
2. OBJETIVOS	57
3. MATERIALES Y MÉTODOS	59
3.1. Materiales.....	59
3.1.1 Material vegetal	59
3.1.1.1 Recolección y almacenamiento	59
3.1.1.2 Obtención de extractos.....	59
3.1.2 Microorganismos.....	60
3.1.3 Medios y condiciones de cultivo.....	62
3.1.3.1 <i>E. coli</i>	62
3.1.3.2 <i>P. pastoris</i>	64
3.1.4 Cebadores	66
3.1.5 Vectores	69
3.1.5.1 pCR®-Blunt II-TOPO®	69
3.1.5.2 pUC19.....	70
3.1.5.3 pET	70
3.1.5.4 pPICZα A.....	72
3.1.5.5 pGAPZα A	74
3.2 Técnicas de manipulación de ácidos nucleicos.....	75
3.2.1 Extracción y purificación de ácidos nucleicos	75
3.2.1.1 Extracción de ARN desde tejidos vegetales	75

3.2.1.2 Extracción de ARN desde <i>P. pastoris</i>	77
3.2.1.3 Extracción de ADN genómico desde <i>P. pastoris</i>	77
3.2.1.4 Extracción de ADN genómico desde <i>G. verum</i>	77
3.2.1.5 Extracción de ADN plasmídico desde <i>E. coli</i>	78
3.2.1.6 Purificación de ADN desde geles de agarosa	78
3.2.1.7 Purificación de ADN desde reacción enzimática	78
3.2.2 Digestión enzimática del ADN.....	78
3.2.3 Defosforilación del ADN.....	79
3.2.4 Fosforilación del ADN.....	79
3.2.5 Ligación del ADN.....	79
3.2.6 Secuenciación.....	80
3.2.7 Síntesis de ácidos nucleicos.....	80
3.2.7.1 Síntesis de ADN complementario	80
3.2.7.2 Síntesis de ADN mediante PCR	81
3.2.7.3 RT-PCR de las cepas de <i>P. pastoris</i> recombinantes.....	82
3.2.8 Electroforesis	83
3.2.8.1 Electroforesis de ARN	83
3.2.8.2 Electroforesis de ADN	84
3.2.9 Determinación de la concentración de ácidos nucleicos.....	84
3.2.9.1 Cuantificación mediante espectrofotometría.....	84
3.2.9.2 Cuantificación mediante electroforesis	85
3.2.10 Amplificación de genes que codifican para PAs de <i>G. verum</i> y <i>F. carica</i> ..	85
3.2.10.1 PCR mediante cebadores degenerados	85
3.2.10.2 RACE.....	86
3.2.10.3 Amplificación del gen <i>gv1AP</i> desde ADN genómico	90
3.3 Generación de microorganismos genéticamente modificados	90
3.3.1 Obtención de cepas de <i>E. coli</i> recombinantes.....	90
3.3.1.1 Construcción del vector pET30a-progv1AP.....	90
3.3.1.2 Construcción de los vectores pET21a-progv1AP y pET21a-profc1AP ..	91
3.3.1.3 Análisis de las construcciones	92
3.3.1.4 Transformación de <i>E. coli</i>	93

3.3.2 Obtención de cepas de <i>P. pastoris</i> recombinantes	94
3.3.2.1 Construcción de los vectores pPICZ α -progv1AP, pPICZ α -profc1AP y pPICZ α -fc1AP	94
3.3.2.2 Construcción de los vectores pGAPZ α -progv1AP, pGAPZ α -progv2AP, pGAPZ α -profc1AP y pPGAPZ α -fc1AP	95
3.3.2.3 Transformación de <i>P. pastoris</i>	97
3.4 Técnicas de análisis y caracterización de proteínas	98
3.4.1 Electroforesis de proteínas en geles de poliacrilamida	98
3.4.2 Cuantificación de proteínas	99
3.4.3 Western blot	100
3.4.4 Extracción de proteínas	101
3.4.4.1 Extracción de proteínas desde cultivos de <i>E. coli</i>	101
3.4.4.2 Extracción de proteínas desde cultivos de <i>P. pastoris</i>	103
3.4.5 Purificación de las proteasas	103
3.4.5.1 Purificación de las proteasas desde <i>E. coli</i>	103
3.4.5.2 Purificación de la progaline B desde <i>P. pastoris</i>	104
3.4.6 Renaturalización de los cuerpos de inclusión desde <i>E. coli</i>	105
3.4.6.1 Método de dilución rápida	105
3.4.6.2 Método de plegado universal	106
3.4.7 Espectrometría de masas	106
3.4.8 Deglicosilación enzimática	107
3.5 Análisis informáticos	107
3.5.1 Análisis de secuencias de nucleótidos y aminoácidos	107
3.5.2 Análisis estadísticos	109
3.6 Ensayos de actividad enzimática	110
3.6.1 Efecto de inhibidores sobre la actividad enzimática	110
3.6.2 Activación de las proteasas recombinantes	110
3.6.3. Zimograma	110

3.6.4 Actividad coagulante de la leche	111
3.6.4.1 Actividad de los extractos vegetales	111
3.6.4.2 Actividad de las enzimas recombinantes	111
3.6.5 Ensayo con BSA como sustrato.	112
3.6.6 Ensayo con FTC-k-caseína como sustrato	112
3.6.7 Ensayo con azocaseína como sustrato.....	113
3.6.8 Ensayo con α -, β - y κ -caseínas bovinas como sustrato	114
3.6.9 Caracterización de la progaline B	115
3.6.9.1 pH de activación.....	115
3.6.9.2 Temperatura óptima y termoestabilidad.....	115
3.6.9.3 pH óptimo	116
3.7 Elaboración y análisis de quesos producidos con la progaline B	116
3.7.1 Elaboración de queso con leche de vaca.....	117
3.7.2 Elaboración de queso con leche de oveja	118
3.7.3 Análisis de las masas queseras	118
4. RESULTADOS	121
4.1 Estudios de la actividad coagulante de la leche	121
4.1.1 Análisis de la actividad coagulante de <i>G. verum</i>	121
4.1.2 Análisis de la actividad coagulante de <i>F. carica</i>	122
4.2 Búsqueda de genes codificantes de proteasas aspárticas de <i>G. verum</i> y <i>F. carica</i>	123
4.2.1 Obtención de secuencias génicas.....	123
4.2.1.1 Secuencias de los genes <i>gv1AP</i> , <i>gv2AP</i> y <i>fc1AP</i>	123
4.2.1.2 Análisis parcial de los intrones del gen <i>gv1AP</i>	129
4.2.1.3 Localización de los tejidos de expresión del gen <i>fc1AP</i>	133
4.2.2 Análisis bioinformático de las secuencias	134
4.2.2.1 Análisis estructural.....	134

4.2.2.2 Análisis filogenético	148
4.3 Expresión de las proteasas aspárticas en <i>E. coli</i>	152
4.3.1 Expresión de la progaline A en el vector pET30a(+).	152
4.3.1.1 Purificación de los cuerpos de inclusión	155
4.3.1.2 Plegado y activación de la progaline A.	156
4.3.2 Expresión de la progaline A y la profigain A en el vector pET21a(+).	158
4.4 Expresión de las proteasas aspárticas en <i>Pichia pastoris</i>.	161
4.4.1 Expresión empleando el vector pPICZ α	162
4.4.2 Expresión empleando el vector pGAPZ α	167
4.5 Producción y caracterización de la progaline B en <i>P. pastoris</i>	171
4.5.1 Estructura	175
4.5.2 Producción en medio de cultivo YPD.	181
4.5.3 Efecto del pH del medio de cultivo en la estructura y actividad	184
4.5.4 Determinación del pH de activación	187
4.5.5 Efecto de la temperatura y el pH sobre la actividad	190
4.5.6 Hidrólisis de α -, β - y κ -caseínas.	193
4.5.7 Conservación del coagulante	197
4.6 Obtención de una cepa superproductora de progaline B	197
4.7 Evaluación de los quesos elaborados con progaline B.	199
4.7.1 Queso fresco de vaca	200
4.7.2 Queso madurado de oveja	204
5. DISCUSIÓN	207
6. CONCLUSIONES	227
7. BIBLIOGRAFÍA	231

8. PUBLICACIONES.....	283
------------------------------	------------





RESUMEN

Todos los cuajos y coagulantes de la leche utilizados para la producción de quesos a nivel industrial son de origen animal o microbiano y pertenecen al grupo de las proteasas aspárticas. Los únicos coagulantes de la leche de origen vegetal utilizados actualmente son los obtenidos a partir de las flores del cardo y su uso es a nivel artesanal. Los quesos producidos con estas proteasas tienen unas características organolépticas diferentes a los producidos con coagulantes de otros orígenes y son muy apreciados por los consumidores. Sin embargo, presentan una serie de problemas como son su elevado precio debido al coste de producción y dificultad en la estandarización de la calidad, que impiden que el producto se comercialice a nivel industrial.

Nuestro equipo de investigación está buscando fuentes alternativas de enzimas coagulantes de leche procedentes de especies vegetales para su aplicación en la industria láctea.

Antiguamente las especies vegetales objeto de esta tesis eran utilizadas para la producción de quesos; así el nombre científico de la especie *Galium verum* hace referencia a su capacidad coagulante (en griego *galaktos* significa leche) y entre los nombres vulgares por los que es conocida figura el de cuajaleches; la otra especie objeto de estudio, *Ficus carica* (higuera) aparece en la definición de coagulante vegetal, a pesar de lo cual no existe ninguna proteasa aspártica descrita de esta especie.

Los objetivos perseguidos con el desarrollo de este trabajo fueron: i) la identificación de genes que codifican para proteasas aspárticas de las especies vegetales *G. verum* y *F. carica* y la

caracterización de las enzimas que codifican. La elección de las especies vegetales objeto de estudio se realizó en base a un segundo objetivo, ii) la producción en sistemas heterólogos de estas enzimas para ser aplicadas en la elaboración de quesos.

El desarrollo de esta tesis doctoral ha contribuido a incrementar el conocimiento sobre las proteasas aspárticas vegetales: hemos identificado y caracterizado *in silico* la secuencia codificante de dos proteasas aspárticas de *G. verum* y una de *F. carica*, así como dos secuencias parciales de esta última especie.

Por otra parte, mediante técnicas de ingeniería genética hemos obtenido una levadura (*Pichia pastoris*) que tiene integrado en su genoma un gen que codifica para una proteasa aspártica de *G. verum*. La enzima recombinante se obtiene desde el medio de cultivo en su forma activa y el sobrenadante del cultivo de la cepa productora puede ser utilizado directamente como coagulante para la elaboración de quesos. Los quesos de vaca elaborados con este coagulante tienen unas características de textura y sensoriales diferentes a las que presentan los elaborados con la quimosina bovina, enzima tradicionalmente utilizada para la coagulación de la leche.

SUMMARY

Almost all rennets and milk- clotting enzymes used in the cheese making industry are of animal or microbial origin belonging to the aspartic proteases group. The only plant milk-clotting aspartic protease comes from the flowers of *Cynara cardunculus* o *Cynara humilis*. Cheeses made with plant enzymes exhibit different organoleptic properties, as compared with those obtained with animal rennet or microbial enzymes and are more expensive due to the difficulties of processes standardization.

Our research group is regularly involved in the finding of new milk-clotting enzymes to be applied in the dairy industry to generate new products with new organoleptic properties.

In the old times, some plants were traditionally used to milk-clotting and these included: i) *Galium verum*, known as “milk-clotting plant” (in Greek *galaktos* means milk); ii) Fig tree or *Ficus carica* whose exudates were employed also traditionally in the cheese making; despite this, no aspartic protease has been described whatsoever.

The pursued objectives of the present work were: i) Identification of the aspartic proteases encoding genes in *G. verum* and *F. carica* as well the characterization of the encoded enzymes. Selection of the plant species was done based on the second objective, ii) the heterologous production to be applied in the cheese making.

The present PhD thesis has contributed to increase our knowledge of plant aspartic proteases. We have characterized *in silico* two aspartic protease- encoding sequences from *G. verum* and one from *F. carica*, as well as two partial sequences from the second.

At the same time and by means of genetic engineering we have constructed a recombinant yeast (*Pichia pastoris*) able to secrete an aspartic protease from *G. verum*. The enzyme is obtained from the culture medium in its active form and may be directly used in the cheese making. The cheeses elaborated with milk from cow have texture and organoleptic properties different from the cheeses elaborated with bovine chymosine, enzyme widely used throughout for the cheese making. This has opened therefore the possibility of creating new types of cheeses by the use of engineered plant aspartic proteases.



ABREVIATURAS

aa: Aminoácidos

ADN: Ácido desoxirribonucleico

ADNc: Ácido desoxirribonucleico complementario

AMV: Virus de la mieloblastosis aviar

AOX: Alcohol oxidasa

ARN: Ácido ribonucleico

ARNm: Ácido ribonucleico mensajero

ARNt: Ácido ribonucleico de transferencia

ATP: Adenosina trifosfato

β-ME: β-mercaptoetanol

BSA: Albúmina sérica bovina (*Bovine Serum Albumin*)

°C: Símbolo de temperatura en grados Celsius

C/P: Relación entre actividad coagulante y actividad proteolítica

C-terminal: Extremo carboxi-terminal

ca. (circa): Cerca de, alrededor de

CAI: Índice de adaptación de codones (*Codon Adaptation Index*)

CBS: Center for Biological Sequence analysis

CN: Caseínas

CP: Cisteín-proteasas

csp: Cantidad suficiente para

CTAB: Bromuro de cetilmetilamonio

dcha.: Derecha

DEPC: Dietilpirocarbonato

DNasa: Desoxirribonucleasa

dNTPs: Desoxinucleótidos trifosfato

DTT: Dithiothreitol

DO: Densidad óptica

E-64: L-transepoxisuccinil-leucilamida-(4-guanidino)-butano

EC: Extracto crudo

EDTA: Ácido etilendiaminotetraacético

EPP: Extracto parcialmente purificado

fam.: Familia

FPC: Quimosina producida por fermentación (*Fermentation produced chymosin*)

FPLC: Cromatografía líquida rápida de proteínas (*Fast protein liquid chromatography*)

g: Fuerza centrífuga relativa

g: Gramo

h: Horas

HC: Cadena pesada (*Heavy Chain*)

HCl: Ácido clorhídrico

H₂O_{mQ}: Agua Milli-Q® (agua ultrapura)

izq.: Izquierda

kb: Kilobase

kDa: KiloDalton

LB: Medio de cultivo Luria-Bertani

LC: Cadena ligera (*Light Chain*)

IMAC: Cromatografía de afinidad de iones metálicos inmovilizados (*Immobilized Metal Affinity Chromatography*)

IPTG: Isopropil-β-D-1-tiogalactopiranosido

M: Molar

m/z: Relación masa carga

MALDI-TOF: Desorción/ionización mediante láser asistida por matriz acoplada a un analizador TOF (tiempo de vuelo) (*Matrix-assisted laser desorption and ionization time-of-flight*)

MCA: Actividad coagulante de la leche (*Milk Clotting Activity*)

MCS: Sitio de clonación múltiple (*Multiple Cloning Site*)

MEGA: Molecular Evolutionary Genetics Analysis

min: Minuto

mL: Mililitro

mM: Milimolar

MMLV: Virus de la leucemia murina Moloney

MOPS: 3-(N-Morpholino) propanesulfonic acid

MP: Metaloproteasas

Mut⁺: Fenotipo silvestre para el gen AOX1 (*Methanol utilization plus*)

Mut^s: Fenotipo mutante para el gen AOX1 (*Methanol utilization slow*)

MV: Máxima verosimilitud

MW: Peso molecular (*Molecular Weight*)

NaOH: Hidróxido de sodio

NC-IUBMB: Comité de Nomenclatura de Enzimas de la Unión Internacional de Bioquímica y Biología Molecular

NCBI: National Center for Biotechnology Information

n.d.: No determinado

NMWCO: Punto de corte de peso molecular, referido a membranas de ultrafiltración (*Nominal Molecular Weight Cut-Off*)

nt: Nucleótidos

N-terminal: Extremo amino terminal

Ori: Origen de replicación

p: p-valor

p/v: Porcentaje peso-volumen

PAs: Proteasas aspárticas

pb: Pares de bases

PCR: Reacción en cadena de la polimerasa (*Polymerase Chain Reaction*)

PDB: Protein Data Bank

PMSF: Fluoruro de fenilmetilsulfonilo

Pro: Prosegmento

ProSA: Protein Structure Analysis

PS: Péptido señal

PSI: Inserto específico de plantas (*Plant Specyfíc Insert*)

Pu: Enzimas purificadas

PVPP: Poly(vinylpopyrrolidone)

RACE PCR: Amplificación rápida de los extremos de ADNc (*Rapid Amplification of cDNA Ends - Polymerase Chain Reaction*)

RNasa: Ribonucleasa

rpm: Revoluciones por minuto

s: Segundo

SAP: Fosfatasa alcalina de camarón (*Shrimp Alkaline Phosphatase*)

SAPLIP: Proteínas tipo saposina (*Saposin-like proteins*)

SDS-PAGE: Electroforesis en gel de poliacrilamida con dodecilsulfato sódico (*Sodium Dodecyl Dulfate Polyacrylamide Gel Electrophoresis*)

SP: Serín-proteasas

TCA: Ácido tricloroacético (*Trichloroacetic Acid*)

U: Unidad de actividad enzimática

μL: Microlitro

μm: Micrometro

μM: Micromolar

US: Unidades Soxhlet

UTR: Región no traducida (*UnTranslated Region*)

v: Volumen

v/v: Porcentaje volumen-volumen

X-Gal: 5-bromo-4-cloro-3-indolil-β-D-galactopiranosid



ÍNDICE DE TABLAS

Tabla 1. Clasificación de las proteasas según el NC-IUBMB.....	2
Tabla 2. Distribución de las proteasas aspárticas entre organismos de diferentes reinos y virus.....	5
Tabla 3. Especies vegetales cuyas proteasas aspárticas han sido investigadas..	21
Tabla 4. Coagulantes de uso común y enzimas que los componen.....	31
Tabla 5. Especies vegetales con actividad coagulante de la leche.....	40
Tabla 6. Enlaces de las caseínas bovinas hidrolizados en solución por las cardosinas y la quimosina.....	47
Tabla 7. Quesos elaborados con extractos del cardo.....	48
Tabla 8. Proteasas aspárticas vegetales recombinantes analizadas como coagulantes de la leche.....	51
Tabla 9. Cepas de <i>Escherichia coli</i> utilizadas en este estudio.....	60
Tabla 10. Cebadores utilizados para PCRs, RACE PCRs y clonación.....	67
Tabla 11. Códigos de acceso de las secuencias de nucleótidos de <i>G. verum</i> y <i>F. carica</i> depositadas en el GenBank.....	127
Tabla 12. Secuencias UTR's obtenidas desde el ADNc utilizando la técnica RACE PCR.....	129
Tabla 13. Predicción de parámetros fisicoquímicos de las proteasas aspárticas identificadas.....	134
Tabla 14. Predicción de los residuos de cisteína implicados en la formación de puentes disulfuro.....	146
Tabla 15. Ensayo FTC-k-caseína de los sobrenadantes de <i>P. pastoris</i> expresando la progaline B.....	174
Tabla 16. Análisis mediante MALDI-TOF de los polipéptidos de 37 kDa y 30.7 kDa de la progaline B recombinante.....	177
Tabla 17. Péptidos de la progaline B recombinante identificados mediante nano LC-MS/MS presentes en el polipéptido de 37 kDa y ausentes en el de 30.7 kDa.....	179
Tabla 18. Fuerza coagulante de la progaline B y preparaciones comerciales.....	193
Tabla 19. Definición de los parámetros de textura analizados.....	203
Tabla 20. Análisis de perfil de textura de quesos elaborados con leche de	

vaca utilizando como coagulante progaine B o quimosina bovina (Chymax™Plus).....	203
Tabla 21. Composición y contenido energético de los quesos elaborados con leche de vaca utilizando como coagulante progaine B o quimosina bovina (Chymax™Plus).....	204
Tabla 22. Análisis de perfil de textura de queso elaborado con leche de oveja utilizando como coagulante progaine B.....	206
Tabla 23. Composición y contenido energético del queso elaborado con leche de oveja utilizando como coagulante progaine B.....	206



ÍNDICE DE FIGURAS

Figura 1. Representación esquemática de la estructura primaria de las proteasas aspárticas de plantas.....	7
Figura 2. Homología del fragmento específico de plantas (PSI) con el dominio “saposina tipo B”.....	9
Figura 3. Estructura cristalográfica de la fitepsina y la cardosina A.....	15
Figura 4. Esquema del procesamiento de la procarnosina A y de la fitepsina...	18
Figura 5. Coagulación enzimática de la leche.....	29
Figura 6. Esquema de los procesos proteolíticos durante la maduración del queso.....	37
Figura 7. <i>Galium verum</i>	54
Figura 8. <i>Ficus carica</i>	56
Figura 9. Dibujo esquemático del vector pCR®-Blunt II-TOPO®.....	69
Figura 10. Dibujo esquemático del vector pUC19.....	70
Figura 11. Esquema del sistema de expresión pET.....	71
Figura 12. Dibujo esquemático del vector pET30a(+).....	71
Figura 13. Dibujo esquemático del vector pET21a(+).....	72
Figura 14. Dibujo esquemático del vector pPICZα A.....	73
Figura 15. Dibujo esquemático del vector pGAPZα A.....	74
Figura 16. Esquema de la técnica RACE PCR para la obtención de la secuencia codificante del gen <i>gv1AP</i>	89
Figura 17. Análisis de la capacidad coagulante de la leche de extractos vegetales de la planta <i>G. verum</i> desecada.....	122
Figura 18. Alineamiento utilizado para el diseño de los cebadores degenerados pAPF y pAPR.....	124
Figura 19. Electroforesis en gel de agarosa de ARN.....	125
Figura 20. Amplificación de secuencias internas que codifican para proteasas aspárticas de <i>G. verum</i> (A) y <i>F. carica</i> (B).....	126
Figura 21. RACE PCR realizada a partir de ADNc de <i>G. verum</i> y <i>F. carica</i>	127
Figura 22. Secuencia del ADN genómico de <i>G. verum</i> que codifica para el gen <i>gv1AP</i>	131

Figura 23. Comparación de la estructura exón-intrón del extremo 3' del gen <i>gv1AP</i> con genes de otras PAs vegetales típicas.....	132
Figura 24. Análisis de la transcripción del gen <i>fc1AP</i> en diferentes tejidos.....	133
Figura 25. Alineamiento de la secuencia de aminoácidos de los genes obtenidos en este estudio con otras PAs vegetales.....	135
Figura 26. Representación gráfica de la estructura primaria de las PAs identificadas en este estudio.....	137
Figura 27. Alineamiento de proteínas de la familia SAPLIP (Saposin A-D y NK-lysin) y los extremos C- y N- terminal del PSI.....	139
Figura 28. Análisis de las secuencias de aminoácidos de la preprogaline A (A), preprogaline B (B) y preprofigain A (C).....	144
Figura 29. Modelo tridimensional de la preprogaline B.....	145
Figura 30. Resultados de validación del modelo 3D de la progaline B mediante el programa ProSA.....	147
Figura 31. Árbol filogenético de la preprogaline A, preprogaline B y preprofigain A y otras PAs vegetales típicas.....	151
Figura 32. Secuencia de aminoácidos de la proteína producto de la expresión de la construcción pET30a-progv1AP.....	152
Figura 33. Análisis mediante SDS-PAGE de la expresión de la progaline A en <i>E. coli</i> utilizando el vector pET30a(+).	153
Figura 34. Influencia de la temperatura de inducción en la producción de la progaline A en <i>E. coli</i> BL21(DE3).....	155
Figura 35. SDS-PAGE para establecer la concentración de imidazol empleada para la purificación mediante IMAC de cuerpos de inclusión producidos en <i>E. coli</i> BL21(DE3) transformada con la construcción pET30a-progv1AP.....	156
Figura 36. SDS-PAGE de cuerpos de inclusión de la progaline A plegados mediante el método de dilución rápida.....	157
Figura 37. Análisis de la actividad proteolítica de la progaline A replegada mediante el método de dilución rápida a temperatura ambiente.....	158
Figura 38. Análisis de la expresión de la profigain A a diferentes temperaturas de cultivo.....	160
Figura 39. Secuencia de las proteasas recombinantes generadas utilizando el vector pPICZα.....	162
Figura 40. Análisis mediante PCR del ADN genómico de los clones de <i>P. pastoris</i> transformada con la construcción pPICZα-progv1AP.....	163

Figura 41. Análisis del fenotipo Mut de las cepas de <i>P. pastoris</i> transformadas con las construcciones pPICZα-progv1AP, pPICZα-profc1AP y pPICZα-fc1AP.....	165
Figura 42. Análisis de la expresión de la preprofigain A, figain A y progaine A en <i>P. pastoris</i> utilizando el vector pPICZα.....	166
Figura 43. Análisis de la transcripción de los genes <i>fc1AP</i> y <i>gv1AP</i> de cepas de <i>P. pastoris</i> transformadas con las construcciones realizadas con el vector pPICZα.....	167
Figura 44. Secuencia de la preprogaine B recombinante generada utilizando el vector pGAPZα.....	168
Figura 45. Construcciones derivadas del vector pGAPZα linearizadas con la enzima BspHI.....	168
Figura 46. Análisis mediante PCR del ADN genómico de las cepas de <i>P. pastoris</i> transformadas con las construcciones (A) pGAPZα-progv1AP (calles 1-4), pGAPZα-progv2AP (calles 5-8) y (B) pGAPZα-profc1AP (calles 1-4) y pPGAPZα-fc1AP (calles 5-8).....	169
Figura 47. Análisis de la expresión de la progaine B, progaine A, profigain A y figain A en <i>P. pastoris</i> utilizando el vector pGAPZα.....	170
Figura 48. RT-PCR de las cepas transformadas con las construcciones pGAPZα-progv1AP (calle 2) y pGAPZα-profc1AP (calle 5).....	171
Figura 49. Análisis mediante SDS-PAGE de los clones de <i>P. pastoris</i> transformados con la construcción pGAPZα-progv2AP.....	172
Figura 50. Actividad coagulante de la leche de los sobrenadantes de <i>P. pastoris</i> transformada con el gen de la progaine B.....	173
Figura 51. Zimograma del sobrenadante de cultivo obtenido desde el clon 7 de <i>P. pastoris</i> expresando la progaine B.....	174
Figura 52. Localización de la secuencia de la preprogaine B utilizada para la producción de anticuerpos empleados en el Western Blot.....	175
Figura 53. Análisis mediante Western blot del sobrenadante de la cepa de <i>P. pastoris</i> PROGVB-7 transcurridas 72 h de cultivo.....	176
Figura 54. Secuencia de aminoácidos de la progaine B recombinante mostrando los péptidos tripticos identificados mediante MALDI-TOF.....	178
Figura 55. Identificación de la secuencia INRLYGNGSSDSIIALK mediante espectrometría de masas en tándem (MS/MS) a partir del ion con una relación m/z 689.67.....	179
Figura 56. Análisis de N-glicosilación de la progaine B.....	180
Figura 57. SDS-PAGE de la progaine B purificada.....	181

Figura 58. Curva de crecimiento de la levadura recombinante y análisis de la progaline B secretada.....	183
Figura 59. Análisis de la influencia del pH del medio de cultivo en el procesamiento de la progaline B secretada por <i>P. pastoris</i>	186
Figura 60. Imagen ampliada del análisis mediante western blot de la progaline B obtenida desde el sobrenadante de un cultivo de <i>P. pastoris</i> en YPD7 bajo condiciones reductoras.....	187
Figura 61. Análisis mediante Western Blot del procesamiento del precursor de la progaline B mediado por pHs.....	189
Figura 62. pH de activación de la progaline B.....	190
Figura 63. Termoeestabilidad de la progaline B.....	191
Figura 64. Efecto de la temperatura sobre la actividad de la progaline B.....	192
Figura 65. Efecto del pH sobre la actividad de la progaline B.....	192
Figura 66. Análisis mediante SDS-PAGE 16.5 % de la hidrólisis de α -, β - y κ -caseínas mediante diferentes coagulantes de la leche (tiempo de incubación 1 hora).....	195
Figura 67. Análisis mediante SDS-PAGE 16.5 % de la hidrólisis de α -, β - y κ -caseínas mediante diferentes coagulantes de la leche (tiempo de incubación 2 horas).....	196
Figura 68. Análisis comparativo de la actividad kappa caseinolítica de los transformantes obtenidos tras un proceso de mejora genética.....	198
Figura 69. Predicción de la frecuencia de uso de codones raros del gen <i>gv2AP</i> cuando se expresa en <i>P. pastoris</i>	199
Figura 70. Imágenes del proceso de elaboración de queso con leche de vaca utilizando como coagulante la progaline B recombinante.....	201
Figura 71. Quesos de leche cruda de vaca elaborados con Chymax® Plus (izq.) y con progaline B recombinante (dcha.).....	202
Figura 72. Queso de oveja elaborado con progaline B transcurridos 50 días de maduración.....	205



INTRODUCCIÓN



1. INTRODUCCIÓN

1.1 Proteasas

Las proteasas son enzimas que hidrolizan enlaces peptídicos de las proteínas, liberando polipéptidos o aminoácidos libres. Las proteasas también son denominadas proteinasas, enzimas proteolíticas o peptidasas, este último término es el recomendado por el Comité de Nomenclatura de Enzimas de la Unión Internacional de Bioquímica y Biología Molecular (NC-IUBMB). Aunque antiguamente estos términos tenían matices en sus significados hoy son considerados sinónimos. Estas enzimas se encuentran en todos los organismos vivos, y están implicadas en un gran número de reacciones fisiológicas, como la digestión de las proteínas de los alimentos, la división celular, la transducción de señales o la apoptosis (Rawlings 2013). Esta función esencial en el metabolismo de los seres vivos también es extrapolable al mundo industrial, desempeñando un papel fundamental en la industria alimentaria, textil, papelera, farmacéutica y biotecnológica. Las proteasas representan cerca del 60 % de las enzimas que se comercializan en el mundo (Adrio y Demain 2014).

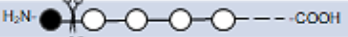
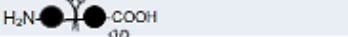


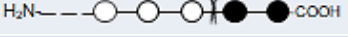

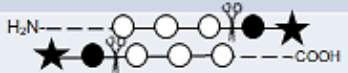
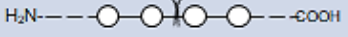
Establecer una clasificación de las proteasas ha sido complicado, debido a la variedad de comportamientos catalíticos que presentan. **Se han utilizado tres criterios diferentes:**

(i) el mecanismo catalítico (Hartley 1960), según esta clasificación las proteasas fueron divididas en cuatro grupos: serín-proteasas, thiol proteasas (actualmente cisteín-proteasas), proteasas ácidas (también conocidas como proteasas aspárticas) y metaloproteasas,

(ii) la reacción catalizada, establecida por la NC-IUBMB (<http://www.sbc.s.qmul.ac.uk/iubmb/>), que clasifica a las proteasas dentro del grupo 3 (hidrolasas) y subgrupo 4 (hidrolasas de enlaces peptídicos). Dentro de esta categoría (EC 3.4) se encuentran las

exopeptidasas y las endopeptidasas. Las exopeptidasas hidrolizan secuencialmente uno o más aminoácidos desde los extremos carboxilo (carboxipeptidasas), amino (aminopeptidasas) o extremos sustituidos, ciclados o unidos por uniones isopeptídicas (omegapeptidasas); las endopeptidasas hidrolizan enlaces peptídicos internos de la proteína (Tabla 1).

Tabla 1. Clasificación de las proteasas según el NC-IUBMB.

	EC 3.4	Peptidasas	Modo de acción
Exopeptidasas	EC 3.4.11	Aminopeptidasas	
	EC 3.4.13	Dipeptidasas	
	EC 3.4.14	Di-peptidil-peptidasas y	
		Tri-peptidil-peptidasas	
	EC 3.4.15	Peptidil-dipeptidasas	
	EC 3.4.16	Carboxipeptidasas tipo serina	
	EC 3.4.17	Carboxipeptidasas tipo metalo	
	EC 3.4.18	Carboxipeptidasas tipo cisteína	
	EC 3.4.19	Omega peptidasas	
Endopeptidasas	EC 3.4.21	Serina endopeptidasas	
	EC 3.4.22	Cisteína endopeptidasas	
	EC 3.4.23	Endopeptidasas aspárticas	
	EC 3.4.24	Metaloenopeptidasas	
	EC 3.4.25	Treonina endopeptidasas	
	EC 3.4.99	Mecanismo desconocido	

Los círculos blancos representan residuos de la cadena polipeptídica; los círculos negros representan bloques de uno, dos o tres residuos terminales escindidos por las proteasas indicadas; las estrellas representan residuos terminales que son sustituidos, ciclados o unidos por enlaces isopeptídicos, sustratos de algunas omega peptidasas. Las tijeras indican la posición del enlace peptídico hidrolizado. Encuadradas aparecen las endopeptidasas aspárticas, objeto de este estudio.

(iii) **la estructura primaria y terciaria**, Rawlings y Barrett (1993) clasificaron las proteasas y sus inhibidores en familias y clanes basándose en este criterio. Las familias agrupan miembros que derivarían de una misma proteína ancestral, con una homología de la secuencia del centro activo estadísticamente significativa. Cada

familia se nombra con una letra denotando el tipo catalítico [A, C, G, M, N, P, S, T y U para aspártica, cisteína, glutámica, metalo, asparragina, mixta, serina, treonina o desconocida (unknown) respectivamente] seguida de un número asignado arbitrariamente, algunas familias son divididas en subfamilias debido a que existe una evidencia de una antigua divergencia. En cada familia una peptidasa que ha sido bien caracterizada es establecida como “*peptidasa tipo*”. El clan es el término utilizado para describir a un conjunto de familias cuyas peptidasas tienen un origen evolutivo común, que han divergido tanto que ya no puede demostrarse su relación mediante comparación de la estructura primaria, pero sí por las similitudes de sus estructuras terciarias. Para un clan, el identificador consta de dos letras, la primera indica el tipo catalítico de las familias incluidas en el clan (con la letra P se designa un clan que contiene familias de más de uno de los tipos catalíticos y con la I inhibidores que son proteínas), seguido de una segunda letra mayúscula asignada arbitrariamente. Algunos clanes se dividen en subclanes en base a evidencias de una antigua divergencia dentro del clan. Este sistema de clasificación dio lugar a la base de datos MEROPS (<https://www.ebi.ac.uk/merops/>), donde se encuentran 912290 proteasas y sus inhibidores, junto con referencias bibliográficas sobre cada una de ellas (Rawlings et al. 2016).

1.2 Proteasas aspárticas (PAs) de plantas

1.2.1 Clasificación

Las proteasas aspárticas (EC 3.4.23) se encuentran ampliamente distribuidas entre organismos de diferentes reinos, así como en virus. Algunas de las enzimas mejor caracterizadas pertenecientes a este grupo son la pepsina y quimosina, presentes en estómagos animales; la cathepsina B, proteasa aspártica lisosomal; la renina, secretada por células del riñón o la proteasa VIH 1 del virus de inmunodeficiencia humana VIH-1 (Szecsi 1992). Las pertenecientes

al reino Plantae se distribuyen principalmente entre las familias A1 (tipo pepsina), A2 (tipo retropepsina), A11 (tipo copia transposón peptidasa), A28 (tipo Ddi) del clan AA, y la familia A22 (tipo preselina-1) del clan AD de la base de datos MEROPS (Tabla 2) (Rawlings et al. 2016). En el año 1997 la NC-IUBMB introdujo el nombre de fitepsina (*L. phyto*, planta) para denominar a todas las proteasas aspárticas típicas procedentes de plantas (Barret 1997), sin embargo, este término no suele emplearse en los artículos científicos y se aplica únicamente a la PA procedente de la cebada (*Hordeum vulgare*) (Kervinen y Wlodawer, 2013)

Faro y Gal (2005), en base a un análisis *in silico* de 50 genes de *Arabidopsis thaliana* que codifican para PAs, establecieron una clasificación de las mismas basadas en los dominios estructurales y en los motivos de la secuencia del sitio activo, que constaba de tres categorías:

(i) **PAs de plantas típicas** (grupo A), cuya estructura es similar a la que presentan las PAs de mamíferos, salvo por la presencia de un fragmento denominado inserto específico de plantas (PSI), que se encuentra entre el extremo N- y C- terminal de la enzima madura. Se dividen en dos subgrupos, A1 y A2.

(ii) **PAs de plantas tipo nucelina** (grupo B), codificadas por genes con delecciones significativas en la región del prosegmento y en el dominio PSI (Chen y Foolad 1997; Takahashi et al. 2008).

(iii) **PAs de plantas atípicas** (grupo C), con características intermedias entre las secuencias de las PAs típicas y las de tipo nucelina. Se dividen en 6 subgrupos C1-C5 y CX.

La mayoría de las PAs de plantas típicas (grupo A) pertenecería a la subfamilia A1A, mientras que los otros dos grupos (grupo B y C) se englobarían casi exclusivamente en la subfamilia A1B de la base de datos MEROPS (Vairo-Cavalli et al. 2013).

Los criterios utilizados para agrupar las PAs codificadas por los genes de *A. thaliana* podrían ser utilizados para la clasificación de PAs procedentes de otras especies vegetales.

Tabla 2. Distribución de las proteasas aspárticas entre organismos de diferentes reinos y virus.

Clan: AA							
Familia	Bacteria	Archaea	Protozoa	Fungi	Plantae	Animalia	Virus
A1	67	2	583	5135	5971	4584	-
A2	13	-	29	32	282	1302	150
A3	-	-	-	-	15	2	43
A9	-	-	-	-	27	21	14
A11	9	-	64	556	2931	480	2
A28	156	5	224	376	635	1462	-
A32	623	1	-	1	1	2	-
Clan: AC							
Familia	Bacteria	Archaea	Protozoa	Fungi	Plantae	Animalia	Virus
A8	6471	6	3	-	1	6	-
Clan: AD							
Familia	Bacteria	Archaea	Protozoa	Fungi	Plantae	Animalia	Virus
A22	9	95	157	310	780	2226	1
A24	6066	86	-	-	3	-	-
Clan: AE							
Familia	Bacteria	Archaea	Protozoa	Fungi	Plantae	Animalia	Virus
A25	323	-	-	-	-	-	-
A31	1795	250	-	-	1	1	-
Clan: AF							
Familia	Bacteria	Archaea	Protozoa	Fungi	Plantae	Animalia	Virus
A26	142	-	-	-	-	-	-
Clan: no asignada							
Familia	Bacteria	Archaea	Protozoa	Fungi	Plantae	Animalia	Virus
A5	-	102	-	-	-	-	-
A36	227	-	-	-	1	-	-
A37	-	59	-	-	-	-	-

*Fuente: MEROPS *the Peptidase Database* consultada el 12/02/2018.

1.2.2 Características comunes de las PAs

La gran mayoría de las PAs de plantas que han sido caracterizadas pertenecen a la familia A1 de la base de datos

MEROPS (Tabla 2) y presentan varias características en común con otros miembros de la misma, como son (Yegin y Dekker 2013):

i) Son sintetizadas como zimógenos, presentando en su estructura primaria un péptido señal y un prosegmento, que preceden a la secuencia de la proteasa madura (Kervinen y Wlodawer 2013).

ii) Poseen una estructura bilobulada, con dos dominios muy similares, sugiriendo que han evolucionado por una duplicación de genes (Tang et al. 1978). Cada dominio contiene uno de los residuos de ácido aspártico que forman parte del sitio activo, estos residuos actúan juntos para favorecer el ataque de una molécula de agua al enlace peptídico (James 2004). En general, estas enzimas actúan mejor sobre enlaces peptídicos que unen aminoácidos con largas cadenas laterales hidrofóbicas, aunque la afinidad por el sustrato varía entre las diferentes proteasas. En la mayoría de las PAs de mamíferos y microbianas cada uno de los residuos de ácido aspártico se encuentra formando parte de la tríada catalítica Asp-Thr-Gly (DTG), mientras que en las plantas uno de los residuos que forman parte del sitio activo se encuentra en la secuencia Asp-Thr-Ser (DSG) (Kervinen y Wlodawer 2013).

iii) Tienen un pH óptimo en el rango ácido, entre 3 y 5, aunque existen algunas excepciones, como la renina, y no necesitan cofactores (Rawlings 2013).

iv) Normalmente son inhibidas por pepstatina A, un hexapéptido producido por *Streptomyces* sp. (Umezawa et al. 1970), la sensibilidad a la pepstatina varía entre diferentes PAs (Kageyama 2002).

1.2.3 Estructura molecular de la PAs de plantas típicas

1.2.3.1 Estructura primaria

La mayoría de las PAs de plantas denominadas por Faro y Gal (2005) como típicas, son sintetizadas como precursores inactivos (monoméricos) y son convertidas posteriormente a enzimas activas mediante un procesamiento proteolítico que dará lugar a enzimas constituidas por dos péptidos no idénticos (heterodiméricas). La secuencia de aminoácidos del precursor es similar en todas ellas y está constituida por cinco regiones (Figura 1): i) un péptido señal, ii) un prosegmento, iii) un dominio N-terminal conteniendo las tríadas catalíticas (DTG y DSG), iv) un inserto específico de plantas (PSI), el cual no está presente en las PAs animales o microbianas y v) un dominio C-terminal (Dunn 2002; Simões y Faro 2004):



Figura 1. Representación esquemática de la estructura primaria de las proteasas aspárticas de plantas. La secuencia incluye un péptido señal (PS), un prosegmento (Pro), un dominio N-terminal conteniendo las tríadas catalíticas (DTG y DSG), un inserto específico de plantas (PSI) y un dominio C-terminal.

i) El **péptido señal hidrofóbico** se encuentra en el extremo N-terminal y es responsable de la translocación de la proteasa al retículo endoplasmático.

ii) El **prosegmento**, constituido por aproximadamente 40 aminoácidos, está presente en las PAs de otros reinos y puede estar implicado en el control de la activación, en el correcto plegamiento y en el direccionamiento celular de la proteasa (Koelsch et al. 1994; Horimoto et al. 2009).

iii) El **dominio N-terminal**, constituido por la cadena polipeptídica mayor o cadena pesada (HC, del inglés *Heavy Chain*), contiene los residuos de ácido aspártico del sitio activo, los cuales forman parte de los motivos DTG y DSG en la mayoría de las PAs de plantas pertenecientes a la familia A1.

iv) Un **inserto específico de plantas**, se emplea la abreviatura PSI para denominarlo (del inglés “*plant-specific insert*”), se encuentra entre los dominios N-terminal y C-terminal, con una longitud aproximada de 100 aminoácidos. El PSI normalmente es procesado y no está presente en la forma activa de la proteasa. Este dominio únicamente ha sido identificado en plantas, y no posee homología con ninguna secuencia de PAs de animales o microorganismos, sin embargo, posee una gran similitud con las saposinas de mamífero (Guruprasad et al. 1994). Las saposinas (SAPs, *Sphingolipid Activator Proteins*) son pequeñas glicoproteínas lisosomales (12-15 kDa), no enzimáticamente activas, que interactúan con lípidos de membrana permitiendo la acción de varias hidrolasas ácidas encargadas de la degradación de esfingolípidos (Kolter y Sandhoff 2010). Las saposinas de mamíferos son sintetizadas como una única proteína precursora (prosaposina) que contiene un péptido señal, dos dominios “saposina tipo A” en los extremos, que son eliminados en la reacción de activación, y cuatro dominios denominados “saposina tipo B” (en la base de datos Prosite figuran como Sap_B, con el código: PS50015), que producirán, mediante procesamiento proteolítico, las saposinas activas A, B, C y D (Figura 2.A). Las proteínas que presentan este tipo de dominio, denominadas SAPLIP (saposin-like proteins), están presentes en una amplia variedad de organismos (la mayoría eucariotas) y tienen en común varias características estructurales, como son: distribución de residuos hidrofóbicos, un posible sitio de N-glicosilación, tres puentes disulfuro intramoleculares, formados entre residuos de cisteína localizados en

posiciones conservadas y una estructura secundaria constituida por cuatro (Munford et al. 1995; Ahn et al. 2006) o cinco hélices α (Liepinsh et al. 1997; Anderson et al. 2003; Hawkins et al. 2005; John et al. 2006) (las hélices 3 y 4 pueden ser consideradas como una única hélice dependiendo del autor) conectadas por bucles cortos (Figura 2.B).

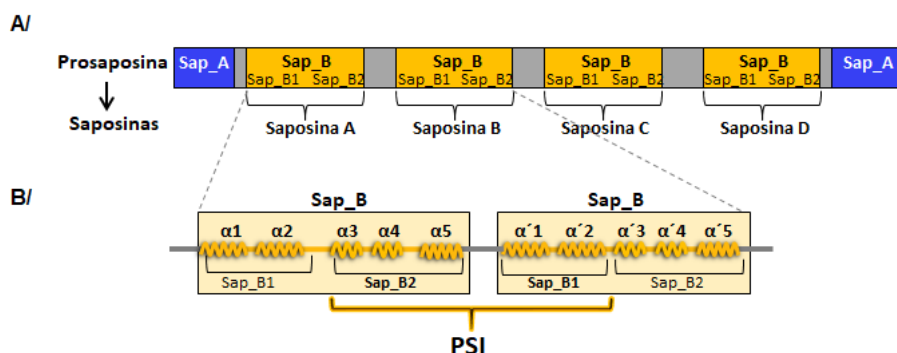


Figura 2. Homología del fragmento específico de plantas (PSI) con el dominio “saposina tipo B”. (A) Estructura de la prosaposina sin el péptido señal. La prosaposina consta de dos dominios “saposina tipo A” (Sap_A) en los extremos (en color azul) y cuatro dominios “saposina tipo B” (Sap_B) (en color naranja) (Prosiste: PDOC50015) organizados como repeticiones en tándem y unidos mediante polylinkers (color gris). El dominio Sap_B puede dividirse en dos subdominios, Sap_B1 (Pfam: PF05184) y Sap_B2 (Pfam: PF03489). La prosaposina da lugar a cuatro proteínas denominadas saposinas A, B, C y D. (B) El dominio PSI (swaposina) es homólogo a la permutación de los extremos N- y C- terminal de un dominio Sap_B (Ponting y Russell 1995; Bliven y Prlić 2012).

La función que desempeñan las SAPLIP es muy variada y está relacionada con la capacidad para interactuar con lípidos de membrana, bien por unión, permeabilización o disrupción de estos. Entre sus roles fisiológicos figuran la participación en el catabolismo de esfingolípidos, presentación de antígenos, actividad antimicrobiana y citolítica, o regulación de la apoptosis (Bruhn 2005).

El PSI presenta las características propias del dominio saposina tipo B (Simões y Faro 2004), pero con una peculiaridad, está constituido por la porción C-terminal de un dominio “saposina tipo B”, denominado en la base de datos Pfam (Finn et al. 2016) como Sap_B2 (código: PF03489) , seguido de una secuencia de unión, y la porción N-terminal del siguiente dominio “saposina tipo B”, denominado en la base de datos Pfam como Sap_B1 (código: PF05184) (Figura 2.B). Debido a que estas estructuras son similares a las saposinas pero con los fragmentos N- y C- terminal intercambiados (Lindqvist y Schneider 1997; Bliven y Prlić 2012) el PSI ha recibido también el nombre de swaposina, procedente de la unión de las palabras *swap* (intercambiar) y saposina (Ponting y Russell 1995).

Se han realizado múltiples estudios para determinar la función del PSI, comprobándose que en algunas proteasas está implicado en la señalización de las PAs a las vacuolas mediante interacción con la membrana (Guruprasad et al. 1994; Kervinen et al. 1999; Tormakangas et al. 2001; Schaaf et al. 2004; Terauchi et al. 2006; Pereira et al. 2013; Pereira et al. 2014); la actividad del PSI de la cipsosina B procedente del cardo es similar a la saposina C, esta actividad se conserva cuando las secuencias N- y C- terminales del PSI son permutadas para que su estructura sea más similar a la de la saposina (Brodelius et al. 2005); se están realizando otros estudios sobre su interacción con la bicapa lipídica y la implicación en sus posibles funciones (Egas et al. 2000; Bryksa et al. 2017); en especies como *Solanum tuberosum* o *Cirsium vulgare* se ha demostrado que el PSI tiene actividad antimicrobiana (Muñoz et al. 2010; Curto et al. 2014; Frey et al. 2018).

La consecuencia de la eliminación del PSI en el plegamiento y actividad proteolítica varía según la procedencia de la proteasa, sugiriendo que juega un papel diferente en cada enzima (Brodelius et al. 2005; Terauchi et al. 2006). Así, el efecto de la eliminación del PSI de la secuencia de diferentes PAs producidas en varios sistemas de

expresión daba lugar a la inactivación de las mismas (White et al. 1999; Lufrano et al. 2012), mientras que en otros casos la proteasa conservaba su actividad proteolítica (Asakura et al. 2000; Tormakangas et al. 2001; Castanheira et al. 2005; Almeida et al. 2014).

v) Dominio C-terminal está constituido por la cadena polipeptídica menor, o cadena ligera (LC, del inglés *Light Chain*). La mayoría de las PAs de plantas presentan cuatro aminoácidos conservados, FAEA en el extremo C- terminal, esta secuencia es suficiente para que se produzca el direccionamiento de una proteína quimérica a las vacuolas según un estudio realizado por Dombrowski y colaboradores (1993), en este trabajo también se determinó la capacidad del PSI para el direccionamiento vacuolar, ambos determinantes de clasificación vacuolar mediaban rutas diferentes (Pereira et al. 2013).

La estructura descrita hasta aquí es la que presentan las PAs de plantas consideradas como típicas, sin embargo existen excepciones a la misma, como son las estructuras presentadas por: la proteasa denominada nucelina, expresada en la nucela del óvulo de cebada (Chen y Foolad 1997) y su proteasa homóloga OSASP1 de *Oryza sativa* (Bi et al. 2005), carentes del dominio PSI; una proteasa CND41 (*Chloroplast Nucleoid DNA binding protein*) codificada por el genoma del cloroplasto de tabaco (Nakano et al. 1997; Murakami et al. 2000), la proteasa CDR1 (*Constitutive Disease Resistance-I*), una proteasa apoplástica de *Arabidopsis* (Xia et al. 2004; Simões et al. 2007), la PCS1 (*Promotion of Cell Survival I*) de *A. thaliana* (Ge et al. 2005) o la nepenthesin-1, procedente de la planta carnívora *Nepenthes gracili* (Kadek et al. 2014). El análisis *in silico* del genoma de *A. thaliana* (Faro y Gal 2005) ha puesto de manifiesto que probablemente las PAs de plantas presentan múltiples variaciones a la

estructura típica de las mayoría de las PAs investigadas hasta el momento.

1.2.3.2 Estructura terciaria

La estructura tridimensional de un gran número de proteasas aspárticas pertenecientes a la familia A1 (base de datos MEROPS) ha sido determinada, mostrando una gran similitud entre proteasas procedentes de diferentes organismos (Dunn 2002). La determinación de la estructura tridimensional de las PAs de origen vegetal ha sido complicada, debido a la dificultad para purificarlas desde el material vegetal, y a que una vez purificadas las preparaciones suelen contener las enzimas en diferentes formas intermedias del proceso de activación. A día de hoy, solo se ha determinado la estructura cristalográfica de la forma inmadura sin el péptido señal de la fitepsina procedente de la cebada (*H. vulgare*) (Kervinen et al. 1999), obtenida mediante el sistema de expresión células de insecto-baculovirus, y de la forma procesada de la cardosina A nativa procedente del cardo (Frazão et al. 1999), cuyos códigos de acceso PDB son 1QDM y 1B5F respectivamente (Figura 3).

La estructura de la fitepsina muestra dos dominios similares entre ellos, formando dos lóbulos, con una estructura secundaria fundamentalmente de tipo barril β . Los dos residuos catalíticos Asp36 y Asp223 pertenecientes a las tríadas DTG y DSG respectivamente (numeración referida a la enzima madura) se encuentran localizados en una gran hendidura situada entre ambos dominios. La estructura está estabilizada mediante tres puentes disulfuro y las cadenas de ambos lóbulos se mantienen unidas mediante interacciones hidrofóbicas y puentes de hidrógeno. El prosegmento se encontraría en la hendidura formada entre los dos dominios, bloqueando parcialmente el sitio activo. El prosegmento y la parte correspondiente al N-terminal de la proteasa madura estarían anclados al sitio activo mediante interacciones iónicas establecidas entre los residuos

Lys11/Tyr13 (posición de los residuos de la secuencia sin el prosegmento) y los ácidos aspárticos del sitio activo. Muchas de las PAs de plantas típicas contienen los residuos Lys/Tyr en posición equivalente a los presentes en *H. vulgare*, lo cual podría implicar un mecanismo de inactivación similar (Simões y Faro 2004).

El papel de estos residuos en la estabilización de la estructura del zimógeno sería similar a la desempeñada por los residuos Lys36p-Tyr37p (numeración pepsina) que se encuentra en el prosegmento de PAs gástricas de mamíferos, como la progastricina y el pepsinógeno (Koelsch et al. 1994; Richter et al. 1998; Bernstein y James 1999). Varias PAs de diferentes orígenes, que son activadas a pH ácido, como la pepsina, la catepsina D o la proteinasa A de las vacuolas de levaduras, parece que comparten este mecanismo de bloqueo del sitio activo, por lo que podría usarse como indicador del pH de activación (Koelsch et al. 1994).

El PSI forma un dominio independiente situado sobre la superficie de la proteína, constituido por 5 hélices α unidas mediante tres puentes disulfuro, que como ya se ha indicado anteriormente es similar a la estructura de las saposinas (Kervinen et al. 1999). Entre las hélices 4 y 5 se encontraría el único sitio de glicosilación presente en la fitepsina (Costa et al. 1997). Aunque las regiones del PSI que muestran homología con las “saposin-like proteins” están permutadas, la estructura terciaria de estos dominios es similar (Egas et al. 2000; Tormakangas et al. 2001; Brodelius et al. 2005; Bryksa et al. 2011).

La estructura cristalográfica de la cardosina A (1B5F) ilustra la conformación de una proteasa aspártica vegetal típica totalmente procesada, en la cual el PSI y el prosegmento no están presentes. Esta estructura estaría formada por dos cadenas de 31 y 15 kDa unidas por interacciones hidrofóbicas y puentes de hidrógeno, que presentan 3 puentes disulfuro, dos de ellos dentro de la cadena polipeptídica mayor o cadena pesada (HC), y uno dentro de la cadena polipeptídica menor o cadena ligera (LC). La estructura presenta dos residuos

glicosilados, alejados del sitio activo. Una característica única de la cardosina A es la presencia en su estructura de la secuencia Arg-Gly-Asp (RGD), en la región C-terminal de la cadena pesada, esta es una secuencia de unión de integrinas en tejidos de mamífero, en donde facilita funciones de reconocimiento celular; otra secuencia presente en la estructura es Lys-Gly-Glu (KGE). Ambas se hallan en una localización expuesta en la superficie molecular y se cree que están implicadas en la interacción entre la cardosina A y la fosfolipasa D, cuya acción conjunta podría estar implicada en procesos degenerativos, como respuesta al estrés, senescencia y/o interacción polen-pistilos (Simões et al. 2005). La estructura global es similar a la que presenta la fitepsina conteniendo el prosegmento, lo cual sugiere que la activación de las PAs típicas implica únicamente pequeños cambios conformacionales. (Frazão et al. 1999).

Recientemente, en un estudio destinado a la optimización de la cardosina B, producida en *kluyveromyces lactis* para su uso en la elaboración de queso, se ha obtenido la estructura cristalográfica de esta enzima con el PSI sustituido por tres residuos de glicina (código PDB: 5NFG), comprobándose que la estructura era similar a la determinada para la cardosina A. La sustitución del dominio PSI no afectaba sustancialmente a la secreción de la proteasa ni a su actividad coagulante de la leche (Almeida et al. 2017).

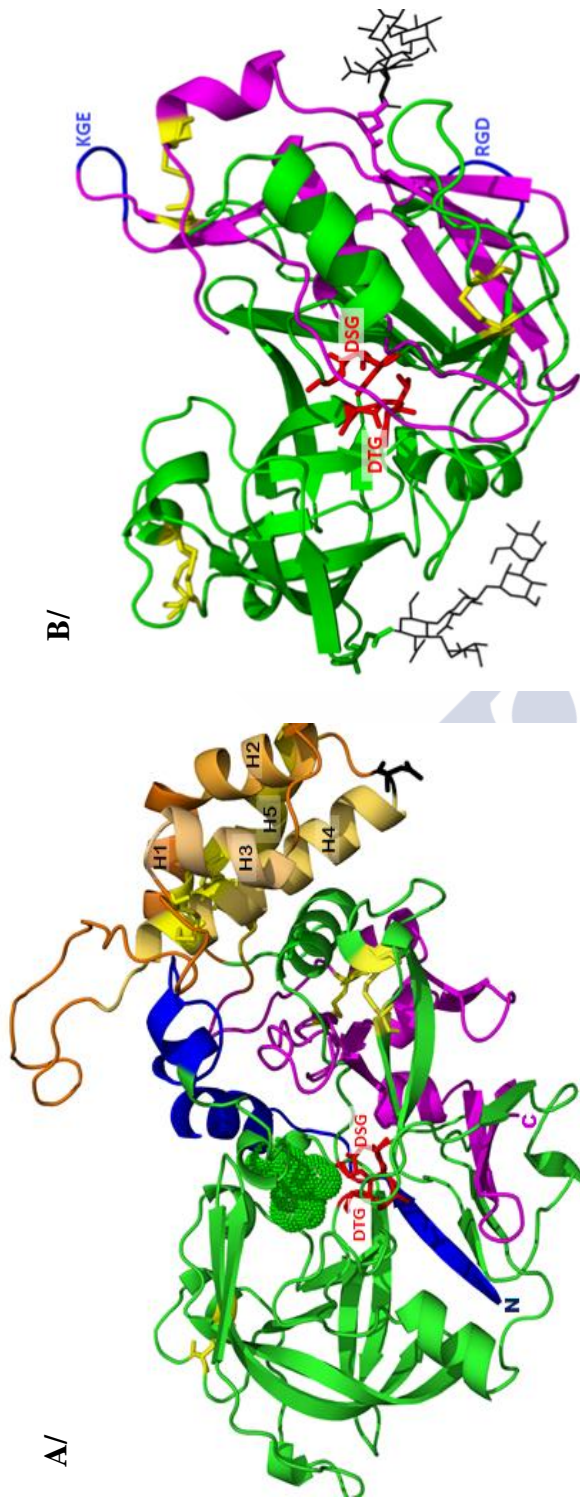


Figura 3. Estructura cristalográfica de la fitepsina y la cardosina. Ambas enzimas son bilobuladas, con dos dominios separados por una hendidura, en la cual se localizan las triadas catalíticas DTG y DSG, que conforman el sitio activo (indicadas como barras rojas) (A) La fitepsina inmadura (modelada con CPHmodels-3.0 utilizando la estructura 1QDM) está constituida por una única cadena en la que se distinguen varias regiones, desde el N al C terminal: el prosegmento (en azul), que junto con los 13 primeros aminoácidos de la enzima madura bloquean el sitio activo (en rojo); los residuos Lys11/Tyr13 de anclaje al sitio activo se indican con esferas verdes; el PSI (tonalidades naranjas) forma un dominio independiente, constituido por cinco α -hélices (H1-H5), y un residuo Asn glicosilado (en negro, la cadena de oligosacáridos no se muestra); las cadenas pesada (HC) y ligera (LC), que constituirán la enzima activa se muestran en verde y magenta respectivamente. Los 6 puentes disulfuro se indican en amarillo (HC:2, LC:1 y PSI:3). (B) La forma activa de la cardosina A (código PDB: 1B5F) está constituida por dos cadenas polipeptídicas, HC (verde) y LC (magenta), ambas glicosiladas (en negro se indican las cadenas de oligosacáridos). Los tres puentes disulfuro se muestran en amarillo (HC:2 y LC:1) y en azul se indican las secuencias correspondientes a los motivos RGD y KGE, característicos de esta proteasa, situados en la superficie molecular. Las representaciones fueron creadas utilizando el software Pymol (DeLano Scientific LLC).

1.2.4 Procesamiento y mecanismos de activación de PAs de plantas típicas.

Las PAs de plantas típicas son sintetizadas como precursores inactivos de cadena simple, y se convierten en enzimas activas mediante un procesamiento que implica varios pasos proteolíticos que incluyen: la eliminación del péptido señal tras la entrada en el retículo endoplasmático, la eliminación del prosegmento, y en la mayoría de los casos la eliminación total o parcial del PSI, formándose así una enzima activa heterodimérica (Glathe et al. 1998; Ramalho-Santos et al. 1998; White et al. 1999; Duarte et al. 2008), esto supone una diferencia con la mayoría de las PAs de otros orígenes, cuyo procesamiento únicamente implica la pérdida del prosegmento, dando lugar a una enzima activa monomérica, con un tamaño molecular aproximado de 35 kDa (Yegin y Dekker 2013). La eliminación del prosegmento en las proteasas aspárticas puede ser desencadenado total o parcialmente por moléculas accesorias o puede tratarse de un proceso autocatalítico que requiera una bajada de pH, como ocurre con las PAs gástricas (Koelsch et al. 1994). El orden y mecanismos que dan lugar a la pérdida del prosegmento y del PSI varía entre las diferentes PAs de plantas, incluso la activación de una misma proteasa puede variar si el procesamiento se produce *in vivo* o *in vitro*, o si se trata de su forma nativa o recombinante (Figura 4):

- El procesamiento de la **procardosina A** de *Cynara cardunculus* ocurre durante la maduración de las flores, el PSI sufre una rotura proteolítica, probablemente mediante una PA y a continuación ocurre la pérdida del prosegmento; el procesamiento *in vitro*, ocurre a través de una forma intermedia que contiene parte del prosegmento y del PSI, el cual está totalmente ausente en la forma activa, formada por una cadena de 31 kDa y otra de 15 kDa (Ramalho-Santos et al. 1998), este procesamiento es similar al que sufre la enzima cuando se expresa de manera transitoria en *Nicotiana tabacum*

o de manera inducible en *A. thaliana*, detectándose la forma activa únicamente en las vacuolas (Duarte et al. 2008). La activación de la procarnosina A clonada en *Escherichia coli* y replegada a partir de cuerpos de inclusión sigue un esquema diferente (Castanheira et al. 2005): la incubación a pH ácido desencadena en primer lugar la pérdida parcial del prosegmento, generándose una forma intermedia, con el PSI todavía presente en su estructura y con actividad proteolítica, lo cual indica que el PSI no jugaría un papel relevante en el proceso de activación; en segundo lugar se procesaría el PSI, la escisión inicial ocurriría en su zona central y el procesamiento procedería bidireccionalmente hacia la zona N- y C- terminal del mismo, originándose la forma activa final constituida por dos cadenas unidas mediante un puente disulfuro, ambas contendrían una parte del PSI, sugiriendo que en el procesamiento *in vivo* de la cardosina A están implicadas otras proteasas (Figura 4.A). La participación de otras proteasas en el procesamiento *in vivo* de otra PA vegetal, la ciprosina B producida en *Pichia* también ha sido propuesta para explicar las diferencias observadas entre el procesamiento *in vivo* e *in vitro* (White et al. 1999).

- En el caso del procesamiento de la **fitepsina** de *H. vulgare*: *in vivo* el prosegmento es totalmente escindido y se produce el procesamiento de una pequeña región interna del PSI, este procesamiento ocurre probablemente en la vacuola, dando lugar a la formación de un heterodímero de 31 y 15 kDa, conteniendo ambas cadenas parte del PSI, la cadena ligera porta además un residuo N-glicosilado. Estos polipéptidos son procesados para dar lugar a la enzima activa, constituida por dos cadenas de 27 y 9 kDa, únicamente la cadena de 27 kDa mantiene parte del PSI, y en el procesamiento de la cadena ligera ha perdido el residuo glicosilado. *In vitro*, el procesamiento de la fitepsina producida en células de insecto, mediado por una bajada de pH, presenta pequeñas diferencias con lo que ocurre *in vivo* (Glathe et al. 1998) (Figura 4.B).

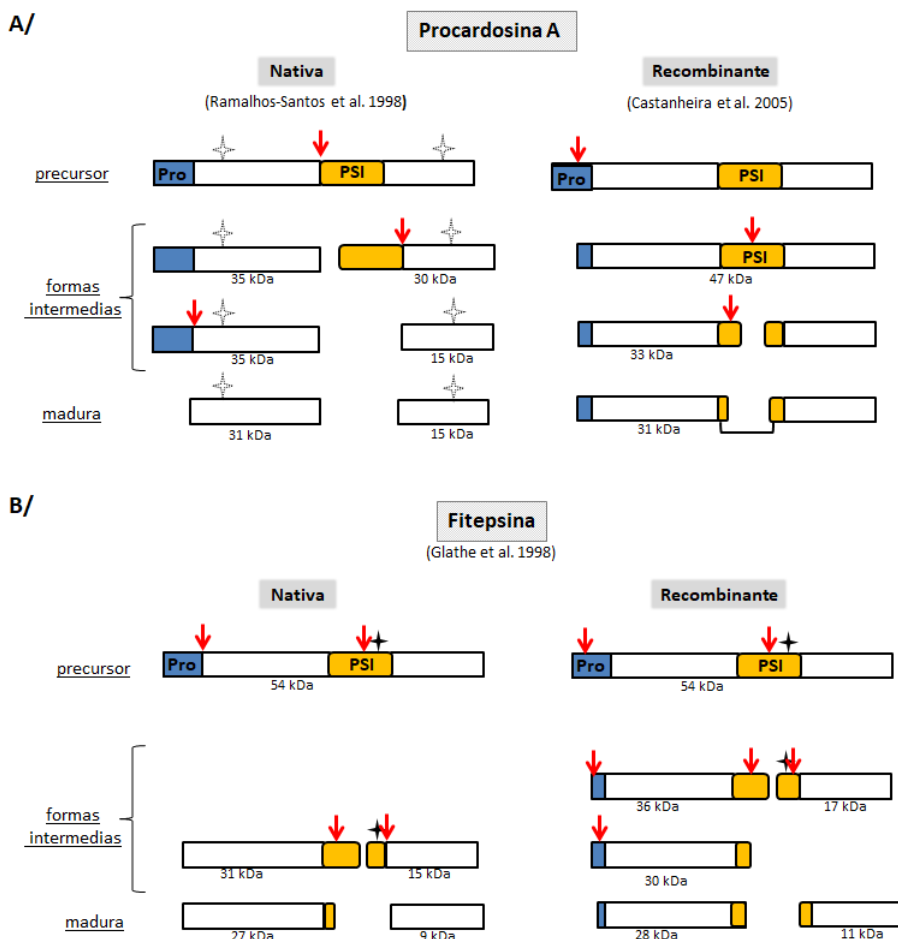


Figura 4. Esquema del procesamiento de la procardosina A y de la fitepsina. (A) Procesamiento de la cardosina A nativa (izq.) (Ramalho-Santos et al. 1998) y producida en *E. coli* (dcha.) (Castanheira et al. 2005). **(B)** Procesamiento de la fitepsina *in vivo* (izq.) e *in vitro* (dcha.) (Glathe et al. 1998). El procesamiento del péptido señal de los precursores de ambas enzimas ocurre en el retículo endoplasmático y se ha omitido. Las flechas rojas indican el sitio donde tiene lugar el procesamiento proteolítico. Las estrellas indican los sitios de glicosilación, en el caso de la procardosina A nativa los autores no hacen referencia a los mismos.

- Durante la activación de la **procardosina B** de *C. cardunculus* expresada de manera transitoria en *N. tabacum* y de manera inducible en *A. thaliana*, sistemas utilizados también para la

expresión de la cardosina A (Duarte et al. 2008), el péptido señal es eliminado tras la entrada de la enzima en el retículo endoplasmático, a continuación ocurre una escisión entre la parte terminal de la cadena pesada y la inicial del PSI, que posteriormente es eliminado por una escisión entre este y la cadena ligera; por último tiene lugar la pérdida del prosegmento, dando lugar a la cardosina B activa, en la cual las cadenas pesada y ligera se mantienen unidas mediante interacciones hidrofóbicas y puentes de hidrógeno (da Costa et al. 2010).

- La **ciprosina B**, con su péptido señal sustituido por el factor- α de *Saccharomyces cerevisiae*, se obtiene desde el hospedador *Pichia pastoris* en su forma activa, sin el prosegmento y con el PSI parcialmente procesado, conservándose la parte inicial y final de este en la cadena pesada y ligera de la enzima madura respectivamente. La escisión del PSI en su parte N-terminal se produce al menos en dos localizaciones diferentes, generando dos cadenas pesadas glicosiladas que difieren en la secuencia de su extremo C-terminal. A diferencia de lo que ocurre con la ciprosina aislada desde las flores, en la ciprosina recombinante las cadenas ligera y pesada se mantienen unidas mediante puentes disulfuro (White et al. 1999). La proteasa aspártica **AtAP1**, de *A. thaliana*, clonada en el mismo sistema de expresión, sufre un procesamiento similar, dando lugar a una enzima activa con una porción del PSI presente en ambas cadenas, que permanecen unidas mediante puentes disulfuro. La cadena ligera portaría en el fragmento del PSI el sitio de glicosilación (Mazorra-Manzano y Yada 2008).

- La forma sin procesar de la **procenprosin** de *Centaurea calcitrapa*, aislada mediante cromatografía desde flores frescas a pH alcalino, se activa mediante incubación a pH ácido, dando lugar a una enzima heterodimérica cuyas cadenas (30 y 16 kDa) son del mismo tamaño que las aisladas desde flores secas, los autores sugieren por ello que *in vivo* tendría lugar una autoactivación proteolítica durante el secado de las flores (Domingos et al. 2000).

- La **procirsin** de *C. vulgare* expresada en *E. coli* sin la parte correspondiente al péptido señal (Lufrano et al. 2012), se produce como cuerpos de inclusión, que una vez plegados dan lugar a una enzima activa monomérica, sin necesidad del procesamiento proteolítico del prosegmento o del PSI, lo cual contrasta con lo descrito para el resto de las PAs estudiadas hasta el momento. Contrariamente a lo que ocurre con la procardosina A, la fitepsina o la procenprosin, la incubación *in vitro* a pH ácido no da lugar a un procesamiento autoproteolítico, los autores argumentan que podrían ser necesaria la acción de otras proteasas para desencadenarlo.

A excepción de este último caso, el procesamiento de la mayoría de las PAs de plantas típicas implica la eliminación del prosegmento y el PSI (total o parcialmente) dando lugar a una enzima heterodimérica activa. Algunos autores de los estudios realizados *in vitro* citados anteriormente, sugieren que además de la rotura autocatalítica, puede ser necesaria la acción de otra/s proteasas para el procesamiento completo de estas enzimas (White et al. 1999; Domingos et al. 2000; Castanheira et al. 2005).

1.2.5 Distribución, localización y funciones biológicas.

Las proteasas aspárticas han sido detectadas, purificadas o clonadas desde especies vegetales muy diversas. En gimnospermas han sido identificadas en semillas de varias especies de pino; en monocotiledóneas se ha determinado su presencia en cebada, arroz, trigo, maíz y en dicotiledóneas en especies como *Arabidopsis*, *Brassica* sp., *Asteraceae* spp., espinacas, judías, tabaco, tomate, girasol, boniato, cacao, pepino, calabaza y ricino; también se han identificado PAs a partir de ADNc del alga unicelular *Chlamydomonas reinhardtii* y del musgo *Physcomitrella patens*. Aunque la gran mayoría han sido aisladas desde semillas también se encuentran en hojas, raíces y flores (Tabla 3).

Tabla 3. *Especies vegetales cuyas proteasas aspárticas han sido investigadas.

Especie vegetal	Fuente	Referencias
Gimnospermas		
<i>Pinus banksiana</i>	Semillas	(Bourgeois y Malek 1991)
<i>Pinus sylvestris</i>	Semillas	(Salmia 1981)
Monocotiledóneas		
<i>Ananas comosus</i>	Pulpa	(Raimbault et al. 2013)
<i>Hordeum vulgare</i> (cebada)	Semillas	(Runeberg-Roos et al. 1991; Chen y Foolad 1997; Glathe et al. 1998)
<i>Hordeum vulgare</i> (cebada)	Flores, tallos, hojas, raíces	(Tormakangas et al. 1994)
<i>Oryza sativa</i> (arroz)	Semillas	(Doi et al. 1980; Asakura et al. 1995; Asakura et al. 1997; Huang et al. 2013)
<i>Triticum aestivum</i> (trigo harinero)	Semillas	(Belozersky et al. 1989; Galleschi y Felicioli 1994; Bleux et al. 1998; Tamura et al. 2007)
<i>Triticum aestivum</i>	Hojas	(Frith et al. 1978)
<i>Zea mays</i> (maíz)	Polen	(Radlowski et al. 1996)
Dicotiledóneas		
<i>Arabidopsis thaliana</i>	Semillas	(D'hondt et al. 1997; Mutlu et al. 1998; Mutlu et al. 1999; Chen et al. 2002; Xia et al. 2004)
<i>Asteraceae spp.</i>	Flores	(Heimgartner et al. 1990; Cordeiro et al. 1994a; Verissimo et al. 1996)
<i>Brassica napus</i>	Semillas	(D'hondt et al. 1997)
<i>Brassica oleraceae</i> (coliflor)	Semillas	(Fujikura y Karssen 1995)

(continuación Tabla 3)

<i>Cucurbita ficifolia</i>	Semillas	(Stachowiak et al. 1994)
<i>Cucurbita maxima</i>	Semillas	(Stachowiak et al. 1994)
<i>Cucumis sativus</i> (pepino)	Semillas	(Wilimowska-Pelc et al. 1983)
<i>Fagopyrum esculentum</i> (trigo sarraceno)	Semillas, hojas, flores tallos, raíces	(Timotijevic et al. 2003; Timotijević et al. 2010)
<i>Ficus racemosa</i>	Latex	(Devaraj et al. 2008)
<i>Glycine max</i> (soja)	Semillas	(Terauchi et al. 2004)
<i>Ipomoea batatas</i> (boniato)	Raíces	(Huang et al. 2009; Chen et al. 2015)
<i>Medicago sativa</i> (alfalfa)	Hojas frescas	(Payie et al. 2000)
<i>Momordica charantia</i> (melón amargo)	Fruto	(Wang et al. 2008)
<i>Nepenthes spp.</i>	Líquido de la jarra	(An et al. 2002; Athauda et al. 2004; Takahashi et al. 2005; Buch et al. 2015)
<i>Nicotiana tabacum</i> (tabaco)	Cloroplasto, hojas	(Nakano et al. 1997; Rodrigo et al. 1991)
<i>Phaseolus vulgaris</i>	Hojas	(Contour-Ansel et al. 2010)
<i>Ricinus communis</i> (ricino)	Semillas	(Hiraiwa et al. 1997)
<i>Solanum tuberosum</i> (patata)	Tubérculo, hojas	(Guevara et al. 1999; Guevara et al. 2001)
<i>Solanum lycopersicum</i> (tomate)	Hojas	(Rodrigo et al. 1989; Schaller y Ryan 1996)
<i>Spinacia oleracea</i> (espinaca)	Hojas	(Kuwabara y Suzuki 1995)
<i>Theobroma cacao</i> (árbol del cacao)	Semillas	(Guilloteau et al. 2005) (Laloi et al. 2002)
<i>Vigna unguiculata</i> (caupí)	Hojas y tallos	(Cruz de Carvalho et al. 2001; Cruz de Carvalho et al. 2004)

(continuación Tabla 3)

Algas verdes		
<i>Chlamydomonas reinhardtii</i>	Cloroplasto	(Almeida et al. 2012)
Musgos		
<i>Physcomitrella patens</i>		(Schaaf et al. 2004)

* Las especies vegetales cuyas PAs han sido investigadas por su actividad coagulante de la leche pueden ser consultadas en la Tabla 5, resaltadas en gris (todas ellas son dictotiledóneas, a excepción de *O. sativa*).

Sumados a los estudios realizados con PAs vegetales obtenidas mediante purificación o clonación, los avances bioinformáticos han permitido realizar estudios *in silico* de estas enzimas, identificándose 69, 96 y 50 genes que probablemente codifican PAs en las especies *A. thaliana* (Faro y Gal 2005; Takahashi et al. 2008), *Vitis vinifera* (Guo et al. 2013) y *O. sativa* (Chen et al. 2009) respectivamente; también se han identificado PAs procedentes de semillas de 33 especies de plantas, pertenecientes a 14 familias diferentes (Amborellaceae, Brassicaceae, Poaceae, Fabaceae, Rutaceae, Cucurbitaceae, Rosaceae, Musaceae, Solanaceae, Salicaceae, Euphorbiaceae, Malvaceae, Poaceae y Vitaceae) (Darabi y Seddigh 2015). El análisis de algunas de estas secuencias utilizando un software que detecta señales de direccionamiento celular, determinó que las proteasas que codifican estarían localizadas en la ruta de secreción, en la mitocondria, unidas a membranas o en el cloroplasto, lo cual indica que deben desempeñar un papel fundamental en numerosos procesos proteolíticos que tienen lugar en diferentes tejidos y localizaciones. Un gran número de las secuencias identificadas *in silico* pertenecen al grupo de las proteasas aspárticas atípicas (Faro y Gal 2005).

Estudios de localización *in situ* han demostrado que la ubicación de las PAs en las plantas no ocurre aleatoriamente, sino que estas se encuentran en tejidos específicos, en función del papel que

desempeñan y de la especie vegetal analizada (Simões y Faro 2004; van der Hoorn 2008).

A nivel celular, las PAs que han sido purificadas hasta el momento estaban localizadas, en su mayoría, en las vacuolas, ejemplos de ello son PAs detectadas en células de las hojas, raíces (Runeberg-Roos et al. 1994) y semillas de la cebada (Marttila et al. 1995), en las semillas de ricino (Hiraiwa et al. 1997), en las papilas estigmáticas de *C. cardunculus* (Ramalho-Santos et al. 1997; Pereira et al. 2013) o en *Arabidopsis* (Mutlu et al. 1999). La mayoría de estas proteasas contenían en su estructura primaria el PSI. Excepciones a esta localización intracelular son las PAs localizadas en la matriz extracelular de tabaco (Rodrigo et al. 1991), de las hojas de tomate (Rodrigo et al. 1989), del tejido floral transmisor de *C. cardunculus* (Vieira et al. 2001; da Costa et al. 2010) y *Arabidopsis* (Xia et al. 2004), las contenidas en el fluido secretado por la planta carnívora *Nepenthes* (An et al. 2002) o una PA de espinaca localizada en los plastidios (Kuwabara y Suzuki 1995).

La función biológica desempeñada por algunas de estas PAs ha sido determinada (Simões y Faro 2004), mientras que en muchas otras únicamente se tienen indicios de su rol fisiológico. Así, en las semillas probablemente desempeñan un papel importante en procesos proteolíticos que ocurren durante su desarrollo y germinación (Runeberg-Roos et al. 1994; Tormakangas et al. 1994; Guilloteau et al. 2005); PAs de *Phaseolus vulgaris* y *Vigna unguiculata* podrían estar implicadas en procesos de degradación inducidos por estrés hídrico (Cruz de Carvalho et al. 2001; Contour-Ansel et al. 2010); la ASPG1 de *A. thaliana* parece que participa en la respuesta al estrés por sequía a través de la vía de señalización dependiente de la fitohormona ABA (ácido abscísico) (Yao et al. 2012); la FeAP9 de *Fagopyrum esculentum* también participaría en respuesta a la sequía y a otras situaciones de estrés abiótico como oscuridad, luz UV-B, heridas o ácido salicílico (Timotijević et al. 2010); varios estudios

sugieren que las PAs de plantas juegan un papel importante en los sistemas de defensa contra el ataque de patógenos y herbívoros (Schaller y Ryan 1996; Guevara et al. 2002; Mendieta et al. 2006; Pagano et al. 2007; Prasad et al. 2009); en plantas carnívoras las PAs podrían estar implicadas en la degradación de insectos para ser utilizados como fuente de nitrógeno (An et al. 2002; Buch et al. 2015); PAs tipo nucelina de *A. thaliana* y del arroz están implicadas en procesos de muerte celular programada, jugando un papel importante en el destino de las células durante la gametogénesis y embriogénesis (Bi et al. 2005; Ge et al. 2005); la SpAP1 del boniato participa en la senescencia foliar mediada por etefón (Chen et al. 2015); otras PAs procedentes de *C. cardunculus* y del arroz participan en el desarrollo floral, reproducción y desarrollo postembrionario (Ramalho-Santos et al. 1997; Radlowski 2005; Pereira et al. 2008; Huang et al. 2013) y la proteasa S5 del arroz forma parte del mecanismo de esterilidad de los híbridos (Chen et al. 2008).

1.2.6 Aplicaciones industriales

El número de proteasas vegetales utilizadas a nivel industrial es pequeño, siendo las más representativas la papaína, bromelaína y ficina, pertenecientes al grupo de las cisteín-proteasas, con aplicación principal en la industria farmacéutica, alimentaria y en la formulación de detergentes (Feijoo-Siota y Villa 2011). Dentro del grupo de las proteasas aspárticas la aplicación más relevante es como coagulantes de la leche en la elaboración de quesos, proceso en el cual se han utilizado desde tiempos remotos y que ha motivado gran parte de las investigaciones desarrolladas sobre estas enzimas (esta aplicación se trata en profundidad en los siguientes apartados por ser objeto de estudio de este trabajo). En la actualidad las PAs vegetales están constatando tener un gran potencial para ser aplicadas en diferentes ámbitos fuera de la industria láctea y dentro de esta para otros fines diferentes al de coagulante, entre ellos fines médicos. Así, proteasas

aspárticas de *C.cardunculus*, en concreto la cardosina A, ha demostrado que tiene una acción específica sobre el colágeno de tipo I, y se ha propuesto que podría ser utilizada en procedimientos médicos de remodelación de la matriz extracelular (Vairo-Cavalli et al. 2013); la ciprosina ha sido objeto de una patente para ser utilizada en composiciones farmacéuticas para su uso como agente antitumoral (Soares et al. 2008); proteasas aspárticas de *S. tuberosum* han sido purificadas y se ha demostrado que tienen actividad antimicrobiana frente a dos patógenos vegetales *Phytophthora infestans* y *Fusarium solani* (Muñoz et al. 2010); el PSI de la procirsina ha sido expresado en la levadura *K. lactis* para investigar su uso como antifúngico (Curto et al. 2014); también se están encontrando nuevas aplicaciones de estas enzimas en la industria alimentaria, así, se ha identificado una PA en semillas de *Theobroma cacao* que está implicada en la producción de péptidos precursores del sabor del cacao (Guilloteau et al. 2005) y dentro del ámbito del sector lácteo se está estudiando su aplicación en la obtención de péptidos bioactivos a partir del suero de la leche (Silva et al. 2006; Cimino et al. 2015) y en la maduración acelerada de los quesos (Prados et al. 2007).

1.3 El queso

El Codex Alimentarius (Codex Standard 283-1978) define el queso como “*el producto blando, semiduro, duro y extraduro, madurado o no madurado, y que puede estar recubierto, en el que la proporción entre las proteínas de suero y la caseína no sea superior a la de la leche, obtenido mediante*”:

(a) ***coagulación total o parcial de la proteína de la leche, leche desnatada/descremada, leche parcialmente desnatada/descremada, nata (crema), nata (crema) de suero o leche de mantequilla/manteca, o de cualquier combinación de estos materiales por acción del cuajo***

u otros coagulantes idóneos, y por escurrimiento parcial del suero que se desprende como consecuencia de dicha coagulación...y/o

(b) técnicas de elaboración que comportan la coagulación de la proteína de la leche y/o de productos obtenidos de la leche que dan un producto final que posee las mismas características físicas, químicas y organolépticas que el producto definido en el apartado (a).

El producto obtenido estará constituido por un entramado de proteínas, las caseínas, que retienen la mayor parte de la materia grasa de la leche y una parte más o menos grande de la fase acuosa, llamada lactosuero. La masa puede ser consumida como tal, bajo la categoría de queso fresco o sufrir una serie de transformaciones que le hacen adquirir caracteres organolépticos específicos, constituyendo el queso maduro.

1.3.1 La leche

La leche es una dispersión coloidal formada por agua, proteínas, grasa, lactosa, minerales, vitaminas y otros componentes minoritarios como, enzimas, urea y sustancias nitrogenadas, su composición varía en función de su procedencia. Las proteínas de la leche se clasifican en dos grupos principales, las caseínas, que son las que constituirán el queso, y las proteínas del suero, que permanecerán solubles durante el proceso de coagulación (Roncada et al. 2012). Las caseínas constituyen aproximadamente el 80 % de las proteínas totales de la leche bovina, se caracterizan por precipitar a pH 4.6 y están fosforiladas. Existen cuatro tipos de caseínas, α_{s1} -, α_{s2} -, β - y κ -caseínas, en una relación aproximada de 4:1:4:1, las tres primeras precipitan en presencia de calcio, sin embargo, la κ -caseína permanece soluble aun en presencia de elevadas concentraciones de este. Estas proteínas se encuentran formando agregados moleculares llamadas micelas, que consisten en un complejo de caseínas unidas mediante fosfato de calcio. La estructura de las micelas ha sido objeto de debate

durante años, y todavía no existe un consenso sobre la misma, pero la mayoría de los modelos coinciden en que las moléculas de κ -caseína se localizan en la superficie de la estructura, estabilizándola (Dalglish y Corredig 2012).

1.3.2 La coagulación de la leche

Como se indica en la definición de queso, la coagulación es el principal fenómeno implicado en la transformación de la leche en queso. Existen tres mecanismos que provocan la modificación de las micelas de caseínas dando lugar a la coagulación de la leche:

i) la acidificación, debida a la adición de ácidos o a la actividad de bacterias ácido-lácticas (**coagulación láctica**), que provocará que las caseínas se vuelvan insolubles cuando se alcanza su punto isoelectrico (pH 4.6) ii) la acción de enzimas denominadas cuajos o coagulantes de la leche (**coagulación enzimática**) o iii) una combinación de los dos procesos anteriores (**coagulación mixta**), que es el tipo de coagulación más común.

La **coagulación enzimática** se divide en dos fases: una fase propiamente enzimática o primaria y una fase de agregación o secundaria (Dalglish 1993). En la **fase primaria**, el cuajo o coagulante, rompe el enlace de la κ -caseína situado entre la fenilalanina 105 (Phe105) y la metionina 106 (Met106), dando lugar a la separación de un fragmento N-terminal, denominado para- κ -caseína (aa 1-105), que es fundamentalmente hidrofóbico y quedará unido a la micela, y el fragmento C-terminal, denominado caseinmacropéptido (aa 106-169), hidrofílico y soluble en agua, por lo que pasará a formar parte del suero, la ruptura de la κ -caseína producirá la desestabilización de la micela. La **fase secundaria** comienza cuando se ha producido la hidrólisis del 80-90 % de la κ -caseína y tiene como resultado la agregación y floculación de las micelas, dando lugar a la formación de un gel lácteo (Guinee y O'Brien 2010). Para que esta fase tenga lugar es necesario que se den unas condiciones de

temperatura, pH y concentración de iones Ca^{+2} apropiadas (Yegin y Dekker 2013) (Figura 5).

La mayor parte del cuajo empleado en el proceso de coagulación se perderá en el suero, pero una pequeña parte quedará retenido en el gel, pudiendo continuar su acción hidrolítica.

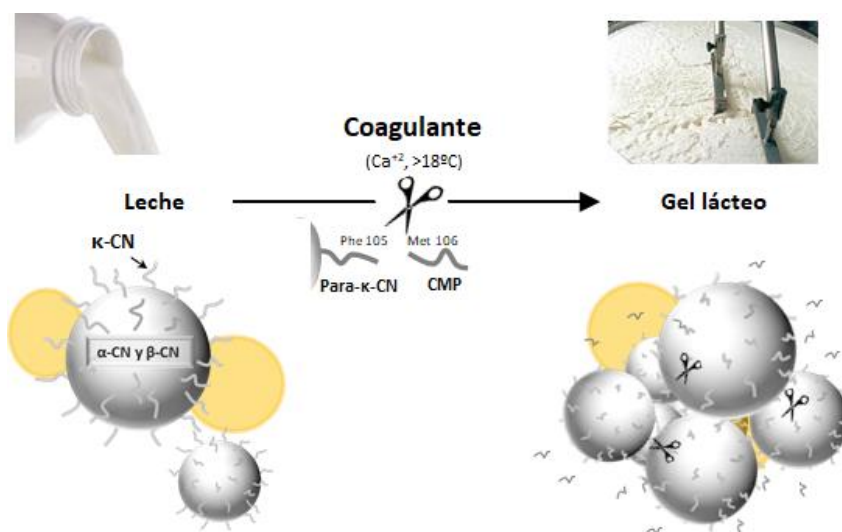


Figura 5. Coagulación enzimática de la leche. Las caseínas (CN) de la leche se encuentran formando micelas (esferas), la κ -caseína (representada por "filamentos") se localiza en la superficie, estabilizando la estructura. El coagulante (representado por unas tijeras) hidroliza la κ -caseína (κ -CN) entre los aminoácidos Phe105-Met106, separando la para- κ -caseína (Para- κ -CN) que permanece unida a la micela y el caseinmacropéptido (CMP), que pasa al suero (fase primaria de la coagulación). Esta modificación provoca la desestabilización de las micelas y su agregación en presencia de calcio y temperatura adecuada, formándose un gel lácteo (fase secundaria). Una parte del coagulante, así como grasa (esferas amarillas) y otros componentes quedarán retenidos en el gel.

1.3.3 Coagulantes de la leche

1.3.3.1 Definición

El Boletín oficial del Estado (BOE-A-1996-4330) define los **coagulantes** como "*preparaciones de proteinasas de origen animal, vegetal o microbiano capaces de provocar la desestabilización de la*

*micela de caseína con formación de un gel lácteo en las condiciones habituales de elaboración del queso.” En esta normativa la palabra **cuajo** queda restringida a “el producto obtenido exclusivamente por extracción de los cuajares de rumiantes cuyo componente activo está constituido por quimosina pura o en mezcla con pepsina de rumiantes”.*

Aunque la coagulación puede ser inducida por varios tipos de proteasas, las enzimas mayoritarias en todos los cuajos y coagulantes, utilizados para la fabricación de quesos a nivel industrial, pertenecen al grupo de las proteasas aspárticas (EC 3.4.23), e hidrolizan el enlace Phe105-Met106 de la κ -caseína bovina, con una excepción, la proteasa de *Cryphonectria parasitica*, que actúa sobre el enlace Ser104-Phe105 (Jacob et al. 2011). El uso de proteasas aspárticas en la elaboración de quesos es una de las primeras aplicaciones conocidas de enzimas en el procesado de alimentos, data aproximadamente de hace 5000 años a.C. Según una leyenda, el queso fue “inventado” por accidente en Oriente Medio, cuando pastores nómada utilizaron los estómagos de rumiantes para transportar la leche por el desierto, la elevada temperatura y las enzimas del interior del odre dieron lugar a la coagulación de la leche (Harboe et al. 2010).

1.3.3.2 Tipos de cuajos y coagulantes

Tradicionalmente los quesos se han elaborado con cuajo animal, pero debido a que la disponibilidad de cuajares de rumiantes no era capaz de cubrir la creciente demanda para la producción de quesos, fue necesaria la búsqueda de fuentes alternativas. Los sustitutos del cuajo animal deberían tener unas cualidades similares a este: una elevada actividad coagulante (especificidad sobre la κ -caseína) en relación con la actividad proteolítica en las condiciones de pH y temperatura utilizadas en la elaboración de los quesos, y una termolabilidad que posibilitase su inactivación en el suero lácteo remanente, para que este pudiese ser utilizado como subproducto.

Las alternativas al uso del cuajo animal utilizadas hoy en día incluyen, enzimas de origen microbiano, proteasas recombinantes producidas mediante microorganismos modificados genéticamente, y en último lugar coagulantes vegetales. Los cuajos/coagulantes pueden ser agrupados así en cuatro categorías (Tabla 4) (Harboe et al. 2010), las cuales comentaremos a grandes rasgos, a excepción de la relativa a los coagulantes vegetales que figuran en un apartado propio (1.4).

Tabla 4. Coagulantes de uso común y enzimas que los componen.

Grupos	Fuentes	Nombres	Componente enzimático activo
1.- Coagulantes animales	Estómago bovino	Cuajo bovino, cuajo de ternero	Quimosina bovina A, B y C, pepsina
		Cuajo en pasta	Igual al anterior y lipasa
	Estómago ovino	Cuajo de cordero, oveja	Quimosina ovina y pepsina
	Estómago caprino	Cuajo de cabrito, cabra	Quimosina caprina y pepsina
2.- FPC*	<i>Aspergillus niger</i>	Quimosina	Quimosina B bovina Quimosina de camello
	<i>Kluyveromyces lactis</i>	Quimosina	Quimosina B
	<i>Escherichia coli</i>	Quimosina	Quimosina A
3.- Coagulantes microbianos	<i>Rhizomucor miehei</i>	Mucorpepsina	Proteasa aspártica de <i>R. miehei</i>
	<i>Rhizomucor pusillus</i>	Mucorpepsina	Proteasa aspártica de <i>R. pusillus</i>
	<i>Cryphonectria parasitica</i>	Endothiapepsina	Proteasa aspártica de <i>C. parasitica</i>
4.- Coagulantes vegetales	<i>Cynara cardunculus</i>	Coagulante vegetal	Ciprosinas 1, 2, y 3 y/o cardosinas A y B

* Quimosina producida por fermentación

1.- Coagulantes animales

El cuajo animal comercial se obtiene a partir del abomaso o cuajar de rumiantes jóvenes, sobre todo terneros. La extracción se realiza a partir de los cuajares secados, salados o congelados que son troceados y macerados en una solución salina durante varios días, transcurridos los cuales el pH se ajusta a 5.0-5.5 y se estandariza. El cuajo contendrá una mezcla de dos enzimas, quimosina y pepsina, en proporciones que dependen de la edad del animal sacrificado y de su tipo de alimentación. Los extractos provenientes de rumiantes jóvenes tienen una mayor proporción de quimosina (80-90 %) con relación a la pepsina, esta proporción se invierte cuando hablamos de animales adultos (Crabbe 2004). La quimosina hidroliza, en su acción inicial, específicamente el caseinmacropéptido desde la κ -caseína durante la fase coagulación (Reid et al. 1997), induciendo la desestabilización de las micelas, mientras que la pepsina es menos específica (Jacob et al. 2011). La quimosina bovina (Uniprot: P00794) es una proteasa aspártica sintetizada como preproquimosina (381 aa), que contiene un péptido señal (16 aa), esencial para que se produzca su secreción al estómago como un precursor inactivo, conocido como proquimosina. En el estómago es activada mediante la eliminación autocatalítica de un prosegmento en la región N-terminal (42 aa) mediada por el pH ácido de este órgano, dando lugar a la enzima activa, con un peso molecular de 35.6 kDa (323 aa) (Kumar et al. 2010). Esta enzima presenta dos variantes (quimosina A y B) que difieren en un único aminoácido. Aunque el cuajo más utilizado es el bovino, también se usan los cuajos procedentes de cordero/oveja y cabrito/cabra, con unas características similares al de origen bovino pero más aptos para coagular leche de su propia especie (Harboe et al. 2010).

2.- Quimosina producida por fermentación (FPC)

La quimosina recombinante, normalmente denominada FPC (*Fermentation Produced Chymosin*) está presente en el mercado desde

1990, siendo la primera enzima recombinante registrada y permitida por la FDA (Administración de Alimentos y Medicamentos de EEUU) para su uso en alimentación (Flamm 1991). Los organismos modificados genéticamente utilizados para la producción de proquimosina de ternera son la cepa de *E. coli* K-12, la levadura *K. lactis* y el hongo *Aspergillus niger* var. *awamori*. Los procesos de obtención de quimosina a partir de cada uno de ellos pueden ser consultados en el Programa Internacional de Seguridad Química (IPCS INCHEM, reports: TRS 806-JECFA 37/12, TRS 806-JECFA 37/13, TRS 806-JECFA 37/14). Un gran número de las investigaciones realizadas sobre coagulantes se centran en la mejora de los sistemas de expresión de la quimosina, así como en la clonación de proteasas homólogas a esta enzima procedentes de otras especies animales, como ciervo, búfalo, jirafa, cerdo, antílope o camello (Vallejo et al. 2008; Kumar et al. 2010; Jacob et al. 2011). La quimosina procedente del camello (Uniprot: Q9GK11) se ha expresado en *A. niger* (Kappeler et al. 2006), comprobándose que es un coagulante más eficiente de la leche de vaca que la propia quimosina bovina, desde el año 2009 está disponible comercialmente bajo el nombre de Chymax-M (Chr. Hansen, Dinamarca).

3.- Coagulantes microbianos

Algunas proteasas extracelulares de origen microbiano se comportan de manera similar a la quimosina, y son adecuadas para la elaboración de quesos. Presentan la ventaja de ser producidas fácilmente mediante fermentación, pero tienen una actividad proteolítica mayor que la quimosina de origen animal y presentan una elevada termoestabilidad. Sumado a esto, los coagulantes microbianos no purificados pueden tener otras actividades secundarias, tipo amilasa, lipasa y celulasa, que pueden afectar negativamente tanto al queso como a la calidad del suero. Los coagulantes comerciales de origen microbiano utilizados industrialmente derivan del medio de

cultivo de los hongos *Rhizomucor miehei*, *Rhizomucor pusillus* y *C. parasitica*, siendo el procedente del primero el más utilizado. Las PAs secretadas por estos hongos se denominan mucorpepsinas (*R. pusillus*, Uniprot: P09177; *R. miehei*, Uniprot: P00799) y endothiapepsinas (Uniprot: P11838), son enzimas monoméricas que al igual que la quimosina se producen como zimógenos, con un péptido señal y un prosegmento, que son eliminados para dar lugar a las enzimas activas. Este tipo de coagulantes también es objeto de numerosas investigaciones orientadas a mejorar sus propiedades coagulantes y a la búsqueda de nuevos microorganismos productores (Feijoo-Siota et al. 2014).

4.- Coagulantes vegetales

Su aplicación está restringida a ámbitos locales, en la elaboración artesanal de quesos. Los coagulantes vegetales se obtienen a partir de frutos, semillas, tallos, raíces, savia, hojas o flores de diversas especies vegetales, normalmente tras un proceso de maceración en agua. El Boletín Oficial del Estado (BOE-A-1996-4330) nombra dos especies en su definición: “*Es el producto de origen vegetal cuyo componente activo tiene actividad coagulante y está constituido por una o varias proteinasas, procedentes de las especies de cardo (Cynara cardunculus, Cynara humilis) e higuera (Ficus carica)*”.

Existen muy pocas regiones donde se elaboren quesos con coagulantes de origen vegetal, estas se restringen a zonas del Mediterráneo, oeste de África y sur de Europa. Los más conocidos son los elaborados, en España y Portugal, utilizando extractos de la flor del cardo, estos quesos presentan unas características muy particulares debido a la elevada actividad proteolítica de las enzimas responsables de la coagulación sobre las caseínas (Roseiro et al. 2003)

1.3.3.3 Consideraciones en la elección de un coagulante

En la búsqueda de nuevas fuentes de coagulantes hay que tener en cuenta una serie de criterios para que puedan ser utilizados por la industria láctea en la elaboración de quesos (Guinee y Wilkinson 1992; Harboe et al. 2010; Fox et al. 2017):

- El coagulante ha de cumplir los requisitos legales en cuanto a pureza, seguridad y ausencia de antibióticos.
- El coste de la producción del coagulante, teniendo en cuenta el rendimiento quesero y el valor del suero lácteo.
- Un elevado ratio actividad coagulante de la leche en relación con la actividad proteolítica (C/P), que evitará una excesiva proteólisis durante la elaboración, y con ello la formación de un gel débil, pérdidas de proteína y grasa en el suero y un bajo rendimiento. Entre los coagulantes utilizados industrialmente este ratio es de mayor a menor: FPC (camello) > FPC (bovina), quimosina bovina > pepsina bovina > coagulantes de *R. miehiei* > coagulante de *C. parasitica* (Harboe et al. 2010).
- La estabilidad del coagulante al pH y temperatura propios del proceso de elaboración del queso.
- Una baja dependencia de la actividad específica del coagulante a variaciones de pH en la región 6.5-6.9. Debido a que el pH de la leche utilizada para la elaboración del queso no es siempre constante, el comportamiento del coagulante no debe tener grandes variaciones ante pequeños cambios de este, para poder controlar el proceso.
- Desarrollo de sabor y textura adecuados.
- Inactivación a la temperatura de pasteurización del suero (72 °C durante 15 s), especialmente para la utilización del suero como subproducto para la elaboración de otros alimentos.
- Efecto del almacenamiento sobre la estabilidad.

1.3.3.4 Participación de los coagulantes en la formación del sabor

La principal función de los coagulantes en la elaboración de los quesos es la digestión de la κ -caseína que desencadena la coagulación de la leche, pero sumada a esta función está el papel que desempeñan durante la maduración de los quesos.

La etapa de maduración es una de las más importantes en el proceso de elaboración de la mayoría de los quesos. De los tres fenómenos bioquímicos que ocurren durante este proceso, glucólisis, lipólisis y proteólisis, este último es el que tiene una mayor importancia, ya que proporciona los sustratos para la formación de sabores por parte de las enzimas procedentes de los microorganismos presentes durante la maduración. La degradación de las proteínas durante la maduración de los quesos es realizada por enzimas procedentes de cinco fuentes: el cuajo o coagulante utilizado, proteasas endógenas de la leche, enzimas de cepas de bacterias lácticas añadidas a la leche (LAB) y desde bacterias lácticas no starter (NLAB) y enzimas desde microorganismos diferentes al cultivo iniciador (Sousa et al. 2001). La proteólisis inicial es llevada a cabo fundamentalmente por el coagulante residual, con excepción de los quesos de pasta cocida, en los que el coagulante puede ser inactivado debido al calentamiento de la cuajada ($\sim 55^\circ\text{C}$) (Sheehan et al. 2007).

Un esquema general del proceso de proteólisis que tiene lugar durante la maduración del queso se muestra en la Figura 6. La hidrólisis de la matriz proteica llevada a cabo por estas enzimas tendrá gran importancia en los cambios de textura del queso (elasticidad, fragilidad, adhesividad, dureza, gomosidad y masticabilidad) y en el desarrollo de sabores y olores, sobre todo en quesos de maduración prolongada.

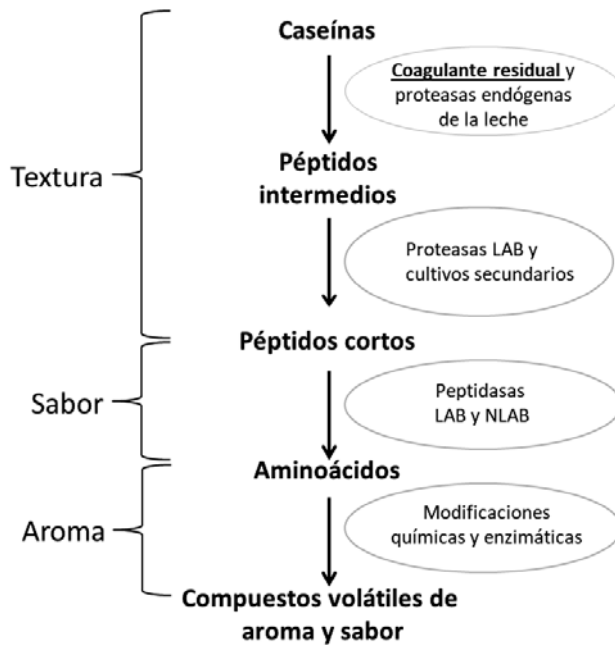


Figura 6. Esquema de los procesos proteolíticos durante la maduración del queso.

A pesar de que solo una pequeña parte del cuajo/coagulante quedará retenido en la cuajada (ca. 6 % de la quimosina o 2 % si se trata de coagulantes microbianos), este tendrá un importante papel en la proteólisis en las primeras etapas de maduración en aquellos quesos cuyo proceso de elaboración no implique su inactivación (Grappin et al. 1985; Fox y McSweeney 1997). Aunque como ya se ha indicado, la quimosina y otros coagulantes tienen una acción específica sobre el enlace Phe105-Met106 de la κ -caseína, en la fase primaria de coagulación, durante el proceso de maduración pueden hidrolizar otros enlaces de las α s1-, α s2-, β - caseínas y para- κ -caseína. Tanto la quimosina como los coagulantes microbianos hidrolizan en esta fase fundamentalmente la α s1 caseína (Upadhyay et al. 2004; Exterkate et al. 1997).

Es por ello importante analizar la degradación de las caseínas realizada por las enzimas investigadas como potenciales coagulantes de la leche, debido a su efecto sobre la consistencia y sabor del queso. Entre las herramientas analíticas utilizadas para realizar esta evaluación se encuentran los geles de poliacrilamida (urea y tricina), HPLC o electroforesis capilar de zona (Sousa et al. 2001).

1.4 Investigaciones sobre coagulantes vegetales

En la obra De Re Rustica, escrita en el año 42 d.C. sobre agronomía (Columela, De Re Rustica), ya se hace referencia al uso de coagulantes vegetales para la elaboración de queso" *[la leche] deberá ser coagulada con cuajo obtenido de cordero o cabrito, pienso que también puede ser coagulada con flores de cardo silvestre (género Cirsium) o semillas de cártamo (Carthamus tinctorius), y también con el líquido que mana de la higuera (género Ficus) cuando se le hace una incisión en la corteza, estando aún verde...*" .

El auge del mundo de la gastronomía, unido al aumento del nicho de mercado de los productos vegetarianos, así como de productos Kosher y Halal, ha llevado a intensificar la búsqueda de nuevos coagulantes vegetales para la producción de quesos con características novedosas. Estas investigaciones se han centrado en extractos vegetales y enzimas contenidas en los mismos expresadas en diferentes hospedadores, tanto eucariotas como procariotas.

1.4.1 Extractos vegetales y sus quesos

Proteasas con actividad coagulante de la leche han sido caracterizadas y/o purificadas desde diversas partes vegetales como, hojas, tallos, frutos, semillas, flores, raíces y látex, de una gran variedad de familias de plantas: Actinidiaceae, Asteraceae, Bromeliaceae, Caricaceae, Apocynaceae, Euphorbiaceae, Moraceae, Rutaceae o Poaceae. En un gran número de estudios se analiza la

actividad coagulante del extracto total, mientras que en otros se realiza una purificación parcial o total de las enzimas contenidas en el mismo para su valoración (Tabla 5). Las enzimas que desencadenan la coagulación de la leche identificadas en estos extractos pertenecen a las familias de las serín-proteasas, cisteín-proteasas y proteasas aspárticas, centrando estas últimas el mayor número de estudios, debido a que sus características catalíticas, similares a la quimosina, las hacen más aptas para este fin. El enlace de la κ -caseína hidrolizado por algunas de estas proteasas no es el que une los aminoácido Phe105-Met106, como ocurre en la mayoría de los coagulantes de la leche, así el extracto de las semillas de *Albizia lebbbeck* hidroliza el enlace Lys116–Thr117 (Egito et al. 2007), las cisteín-proteasas de *Cynanchum otophyllum* el Ser132-Thr133 (Luo et al. 2018), la cisteín-proteasa procerain B hidroliza el enlace Ala90-Gln91 (Zhang et al. 2015) o las cisteín-proteasas de *Zingiber officinale* los enlaces Ala90-Glu91 y His102-Leu103 (Huang et al. 2011).



Tabla 5. Especies vegetales con actividad coagulante de la leche.

*Especie (nombre común)	Tejido	**Enzimas	***Preparación (nombre)	Referencia
<i>Actinidia chinensis</i>	Frutos	n.d.	EPP (actinidin)	(Katsaros et al. 2010; Alirezaei et al. 2011)
<i>Actinidia deliciosa</i> (kiwi)	Frutos	n.d.	EPP (actinidin)	(Lo Piero et al. 2011; Grozdanovic et al. 2013)
<i>Albizia julibrissin</i>	Semillas	CP	Pu	(Otani et al. 1991)
<i>Albizia lebeck</i> (ébano de oriente)	Semillas	n.d.	EC	(Egito et al. 2007)
<i>Aloe variegata</i>	Hojas	n.d.	EC	(Ahmed y Helmy 2012)
<i>Ananas comosus</i> (piña)	Fruto	n.d.	EC	(Soni y Rakha 2017)
<i>Balanites aegyptiaca</i> (datilero del desierto)	Frutos	PA SP	EPP	(Beka et al. 2014)
<i>Bromelia balansae</i>	Frutos	CP	Pu (balansain I)	(Pardo et al. 2001)
<i>Bromelia hieronymi</i>	Frutos	CP	EC (hyeronimain)	(Bruno et al. 2010)
<i>Calotropis gigantea</i> (algodón de seda)	Látex	CP	EC	(Anusha et al. 2013; Anusha et al. 2014)
<i>Calotropis procera</i> (manzana de sodoma)	Látex	C	EC	(Freitas et al. 2016)
	Látex	n.d.	EC	(Oseni y Ekperigin 2013)
	Látex	CP	Pu (procerainB)	(Singh y Dubey 2011)
	Hojas	n.d.	EPP	(Aworh y Muller 1987)

(continuación Tabla 5)

<i>Calotropis procera/ Carica papaya/ Citrus x limon</i>		Hojas/Hojas/Fruto	n. d.	EC	(Akinloye y Adewumi 2014)
<i>Centaurea calcitrapa</i> (cardo estrellado)	Flores		PA	Pu	(Domingos et al. 1998)
	Cultivo celular		PA	Pu	(Reis et al. 2000; Raposo y Domingos 2008)
<i>Citrus aurantium</i> (naranja amargo)	Flores		PA	EC	(Mazorra-Manzano et al. 2013a)
<i>Cryptostegia grandiflora</i>	Látex		CP	EC	(Freitas et al. 2016)
<i>Cucumis melo</i> (melón)	Fruto		SP	Pu (cucumisin)	(Uchikoba y Kaneda 1996; Gagaoua et al. 2017)
<i>Cynanchum otophyllum</i>	Flores		CP	Pu (QA y QC)	(Luo et al. 2018)
<i>Cynara cardunculus</i> (cardo)	Flores		PA	EC o EPP	(Campos et al. 1990; Cordeiro et al. 1994a; Sousa y Malcata 1996; Chen et al. 2003a)
	Flores		PA	Pu (cardosinas A y B)	(Verissimo et al. 1996)
	Flores		PA	Pu (cynarinas 1,2,3)	(Heimgartner et al. 1990)
<i>Cynara humilis</i> (cardo borriquero)	Flores		PA	Pu	(Vairo-Cavalli et al. 2013)
<i>Cynara scolymus</i> (alcachofa)	Flores			EPP	(Nouani et al. 2009; Llorente et al. 2014; Abd El-Sal et al. 2017)
	Flores		PA	Pu (cynarinas A,B,C)	(Llorente et al. 2004; Sidrach et al. 2005; Chazarra et al. 2007)

(continuación Tabla 5)

<i>Dregea sinensis</i>	Tallo	CP	Pu (procerain B)	(Zhang et al. 2015)
<i>Euphorbia</i> spp.	Látex	n.d.	EC	(Mahajan y Chaudhari 2014)
<i>Euphorbia microsciadia</i>	Látex	CP	Pu (microsciadin)	(Rezanejad et al. 2015)
<i>Euphorbia neriifolia</i>	Látex	SP	Pu (neriifolin)	(Yadav et al. 2011; Yadav et al. 2012)
<i>Euphorbia nivulia</i>	Látex	CP	Pu (nivulian-II)	(Badgujar y Mahajan 2014)
<i>Euphorbia amygdaloides</i>	Látex	n.d.	Pu	(Demir et al. 2005)
<i>Ficus carica</i> (higuera)	Látex	n.d.	EPP	(Nouani et al. 2009)
<i>Ficus carica</i>	Látex	n.d.	EC	(Bornaz et al. 2010; Faccia et al. 2012)
<i>Ficus racemosa</i>	Látex	PA	Pu	(Devaraj et al. 2008)
<i>Ficus religiosa</i>	Látex	SP	Pu (religiosin B)	(Kumari et al. 2010; Kumari et al. 2012)
<i>Helianthus annuus</i> (girasol)	Semillas	n.d.	EC EPP	(Egito et al. 2007) (Nasr et al. 2016)
<i>Jacaratia corumbensis</i>	Látex de raíz	CP	EPP	(Duarte et al. 2009)
Kiwi/Melón/Jengibre	Frutos/Frutos/ Rizomas	n.d.	EC	(Mazorra-Manzano et al. 2013b)
<i>Lactuca sativa</i> (lechuga)	Hojas	SP	Pu (lettucine)	(Lo Piero et al. 2002)
<i>Maclura pomifera</i>	Látex	SP	EC	(Corrons et al. 2012)

(continuación Tabla 5)

<i>Moringa oleifera</i> y <i>C. procera</i>	Semillas y Látex	n.d.	EC	(Mahami et al. 2012)
<i>Moringa oleifera</i>	Flores	CP, SP, MP y PA	EC	(Pontual et al. 2012)
<i>Ongokea gore</i>	Corteza	n.d.	EC	(Libouga et al. 2004)
<i>Onopordum acanthium</i> (cardo de algodón)	Flores	PA	EPP (onopordosin)	(Brutti et al. 2012)
<i>Onopordum turcicum</i>	Semillas, hojas y flores	PAs	EPP	(Tamer 1993)
<i>Oryza sativa</i> (arroz)	Semillas	PA	Pu (oryzasin)	(Asakura et al. 1997)
<i>Opuntia</i> sp. (cactus)	Cladodios y frutos	CP y otras	EC	(Teixeira et al. 2000)
<i>Papilionoideae</i> spp.	Raíces, hojas	n.d.	EC	(Lopes et al. 1998)
<i>Primula veris</i>	Planta completa	n.d.	Pu	(Demir et al. 2007)
<i>Sideroxylon obtusifolium</i> (quixaba)	Látex	CP, SP, PA	EC	(Silva et al. 2013)
<i>Silybum marianum</i> (cardo mariano)	Flores	PA	EC	(Vairo-Cavalli et al. 2005; Vairo-Cavalli et al. 2008)
<i>Solanum dubium</i>	Semillas	SPs	EPP (dubiumin)	(Talib et al. 2009; Ahmed et al. 2010; Kheir et al. 2011)
<i>Solanum elaeagnifolium</i>	Fruto y semillas	n.d.	EC	(Chávez-Garay et al. 2016)
	Frutos	n.d.	EC	(Néstor et al. 2012)
<i>Solanum</i> spp.	Frutos	n.d.	EC	(Guíama et al. 2010)

(continuación Tabla 5)

<i>Streblus asper</i>	Látex	SP	Pu (streblin)	(Tripathi et al. 2011)
<i>Streblus asper</i>	Ramas	PA	P	(Senthilkumar et al. 2006)
<i>Vitis vinifera</i>	Frutos	n.d.	EC	(Felix da Silva et al. 2015)
<i>Withania coagulans</i>	Frutos	PA	EPP	(Pezeshki et al. 2011; Kazemipour et al. 2017)
<i>Withania coagulans</i>	Frutos	PA	Pu	(Salehi et al. 2017)
<i>Zingiber officinale</i> (jenjibre)	Rizoma	CPs	Pu (zingibain)	(Huang et al. 2011)
	Rizoma	CP	Pu	(Hashim et al. 2011b)

* Las especies resaltadas en gris presentan proteasas aspárticas responsables de la actividad coagulante en el extracto.

** Familia a la que pertenecen las proteasas analizadas: no determinada (n.d.), proteasas aspárticas (PA), cistein-proteasas (CP), serín-proteasas (SP), metaloproteasas (MP)

*** Preparación: Extracto crudo (EC), extracto parcialmente purificado (EPP), enzimas purificadas (Pu)

Los extractos vegetales con actividad coagulante de la leche mejor caracterizados son los procedentes de varias especies pertenecientes a la familia Asteraceae, debido a su uso en la elaboración de quesos (Pires et al. 1994; Vioque et al. 2000). Las enzimas responsables de la coagulación contenidas en estos extractos pertenecen al grupo de las proteasas aspárticas. Así, a partir de *C. cardunculus* (cardo) se han identificado dos tipos de PAs:

- las **cynarasas 1, 2 y 3** (Heimgartner et al. 1990), denominadas posteriormente cipsosinas, aisladas desde flores secas en condiciones alcalinas (pH 8.3) (Cordeiro et al. 1994b; Cordeiro et al. 1994a), siendo la cipsosina 3, denominada también cipsosina B, la que presenta una mayor similitud con la quimosina bovina (Cordeiro et al. 1992).
- las **cardosinas A y B**, aisladas desde flores frescas en condiciones ácidas (Verissimo et al. 1996; Ramalho-Santos et al. 1998; Faro et al. 1999; Vieira et al. 2001), siendo la cardosina A, en términos de especificidad y parámetros cinéticos similar a la quimosina, mientras que la cardosina B se asemeja más a la pepsina (Verissimo et al. 1995; Silva y Malcata 2000).

Cardosinas y cipsosinas nunca han sido copurificadas. Estas proteasas se encuentran en los pistilos en gran cantidad, lo cual es una característica poco común en el resto de las plantas, donde las enzimas de esta clase son aisladas principalmente desde las semillas u hojas (Tabla 3), y en concentraciones menores. Las cardosinas y cipsosinas forman parte de una familia génica (Pimentel et al. 2007) de la cual se han caracterizado hasta el momento nueve PAs: las cipsosinas 1, 2 y 3 y las cardosinas A, B, E, F, G y H (Sarmiento et al. 2009).

Otras PAs pertenecientes a la familia del cardo han sido purificadas mediante extracción ácida y técnicas cromatográficas a partir de flores frescas de *C. humilis* (Vairo-Cavalli et al. 2013), *Onopordum acanthium* (Brutti et al. 2012) o *Silybum marianum* (Vairo-Cavalli et al. 2005) y desde flores frescas (Llorente et al. 2004)

y secas (Sidrach et al. 2005) de *Cynara scolymus* y de *C. calcitrapa* (Domingos et al. 1998).

Aunque menos estudiadas, también se han identificado proteasas aspárticas con actividad coagulante de la leche en especies de otras familias, como *Citrus aurantium* (fam. Rutaceae) (Mazorra-Manzano et al. 2013a), *F. racemosa* (fam. Moraceae) (Devaraj et al. 2008), *O. sativa* (fam. Poaceae) (Asakura et al. 1997), *Streblus asper* (fam. Moraceae) (Senthilkumar et al. 2006) o *Withania coagulans* (fam. Solanaceae) (Salehi et al. 2017).

La mayoría de los extractos que figuran en la Tabla 5 no se utilizan en la elaboración de quesos, debido a que presentan una elevada actividad proteolítica en relación a su actividad coagulante (Ahmed et al. 2009; Pontual et al. 2012; Mazorra-Manzano et al. 2013b), dando lugar a sabores amargos, a defectos en la textura y/o a un bajo rendimiento (Nuñez et al. 1991; Tejada et al. 2008; Pino et al. 2009; Pezeshki et al. 2011; El-Kholy 2015). Una excepción son los extractos acuosos de las flores del cardo, que se utilizan en España y Portugal para elaborar quesos de oveja y de cabra, muchos de ellos con Denominación de Origen Protegida. Las PAs contenidas en estos extractos, al igual que la quimosina, tienen una elevada actividad coagulante de la leche, hidrolizando en la fase primaria de coagulación el enlace Phe105-Met106 de la κ -caseína bovina y ovina (Macedo et al. 1993), y el enlace Lys116-Thr117 de la caprina (Sousa y Malcata 1998), sin embargo la hidrólisis de las α - y β - caseínas bovinas y ovinas es mayor que la efectuada por la quimosina bovina (Macedo et al. 1993; Macedo et al. 1996; Sanjuán et al. 2002) (Tabla 6). Esta excesiva proteólisis es la responsable de las características únicas de textura, sabor y aroma que caracterizan a los quesos elaborados con estos extractos (Reis y Malcata 2011; Ordiales et al. 2014) (Tabla 7).

Tabla 6. Enlaces de las caseínas bovinas hidrolizados en solución por las cardosinas y la quimosina (Macedo et al. 1996).

	Caseínas bovinas			
	α s1	α s2	β	κ
Quimosina	Phe23-Phe24		Ala189-Tyr190 Leu192-Tyr193	Phe105-Met106
Extracto cardo	Phe23-Phe24 Trp164-Tyr165 Tyr166-Val167 Tyr165-Tyr166 Phe153-Tyr154 Phe145-Tyr146 Leu149-Phe150 Leu156-Asp157 Ala163-Trp164	Phe88-Tyr89 Tyr95-Leu96	Leu192-Tyr193 Leu191-Leu192 Leu165-Ser166 Phe190-Leu191 Ala189-Phe190 Leu127-Thr128	Phe105-Met106

*Enlaces peptídicos en orden decreciente a la susceptibilidad de ser hidrolizados

Los extractos del cardo también son utilizados en Italia para elaborar un queso con leche de vaca denominado Caciotta (Aquilanti et al. 2011) y en Marruecos una variedad llamada Jben (Mennane et al. 2007). Otros coagulantes vegetales utilizados a nivel local son las bayas de la planta *Solanum dubium*, para la producción de un queso tradicional en Sudán (Kheir et al. 2011; Mustafa et al. 2013), las bayas de *Solanum elaeagnifolium*, utilizadas en Méjico para elaborar un queso fresco de pasta hilada (queso asadero) (Néstor et al. 2012) o el zumo extraído de *Calotropis procera* (manzana de sodoma) utilizado en regiones de Nigeria y Benin.

Los quesos elaborados con los extractos procedentes de especies de la familia Asteraceae han centrado la mayor parte de las investigaciones orientadas a determinar los efectos de los coagulantes vegetales sobre las masas queseras. Vieira de Sá y Barbosa (de Sá y Barbosa 1972) fueron los primeros en comparar las características de diferentes tipos de quesos (Edam, Serra y Roquefort) elaborados con extractos de *C. cardunculus* o con cuajo animal, comprobando que el coagulante vegetal era más adecuado para la elaboración de queso de

origen ovino y de cuerpo blando, tipo Serra, pero que debido a su elevada actividad proteolítica presentaba algunos problemas para la elaboración de quesos tipo Edam; en quesos tipo Roquefort daba buenos resultados, aunque con bajo rendimiento.

Tabla 7. Quesos elaborados con extractos del cardo (Roseiro et al. 2003; Reis y Malcata 2011).

País	Queso	Tipo de leche
España	Casar de Cáceres	Oveja
	Torta del Casar	Oveja
	La Serena	Oveja
	Los Pedroches	Oveja
	Flor de Guía	Oveja y cabra
Portugal	Serra da Estrela	Oveja
	Serpa	Oveja
	Castelo Branco	Oveja
	Évora	Oveja
	Azeitão	Oveja



En las últimas décadas se han realizado un gran número de estudios sobre la influencia del uso de extractos del cardo y diferentes parámetros de elaboración en las características finales de diferentes tipos de quesos, como: Picante da Beira Baixa (Freitas y Malcata 1996), Cheddar (O'Mahony et al. 2003), Los Pedroches (Vioque et al. 2000; Sanjuán et al. 2002), diferentes quesos italianos (Barbosa et al. 1981), queso fresco de cabra (García et al. 2012), queso Serra da Estrela (Macedo y Malcata 1996), Torta del Casar (Francisco-José et al. 2010; Ordiales et al. 2012; Ordiales et al. 2014), queso Azeitão (Pinho et al. 2001) y quesos tradicionales portugueses (Pereira et al. 2011).

A pesar de que el uso de los coagulantes vegetales se ha asociado al desarrollo de sabores amargos debido a su actividad proteolítica, varios estudios han determinado que, bajo parámetros de elaboración adecuados, los quesos producidos con estos coagulantes pueden presentar unas características similares a los obtenidos con

quimosina, incluso cuando se utiliza leche de origen bovino. Así, un extracto de *O. acanthium* (fam. Asteraceae) fue utilizado para elaborar un queso semiduro con leche bovina pasteurizada cuyas características organolépticas eran similares a otros quesos comerciales (Brutti et al. 2012); el queso Peshawari elaborado con leche de vaca coagulada con una proteasa procedente del jengibre recibió una mayor puntuación en los atributos organolépticos que la obtenida con cuajo de ternera (Hashim et al. 2011a); Agboola y colaboradores (2004) compararon el contenido en péptidos hidrofóbicos e hidrofílicos en quesos elaborados con leche de oveja utilizando cuajo de ternera, un coagulante microbiano y extractos de cardo, comprobando que este último contenía un mayor número de péptidos de ambas clases, pero un menor número de péptidos asociados al sabor amargo (péptidos hidrofóbicos de masas entre 165–6500 g mol⁻¹) que el coagulante microbiano, además, la intensidad de amargor, percibido por un panel de cata entrenado, era menor que en los quesos elaborados con cuajo y coagulante microbiano, probablemente debido a la modificación del sabor por otros componentes de queso; el queso italiano pecorino elaborado con extracto de cardo resultó ser más suave, cremoso y menos amargo que el preparado con cuajo animal o un coagulante microbiano (Chen et al. 2003a); un extracto parcialmente purificado obtenido a partir de semillas de *Albizia julibrissin* resultó ser apto para la elaboración de queso tipo Gouda, presentando un rendimiento comparable al producido con quimosina y sin desarrollo de sabores amargos transcurridos tres meses de maduración (Otani et al. 1991).

Otras investigaciones se han centrado en el desarrollo de estrategias para solventar algunos de los inconvenientes que presentan los coagulantes vegetales para su aplicación en la industria láctea, entre ellas se encuentran (Ben Amira et al. 2017a):

- ✓ el uso del coagulante vegetal en combinación con cuajo animal;

- ✓ la elección apropiada del tipo de leche; la leche de origen ovino y caprino parecen ser las más adecuadas para los coagulantes vegetales (Tavaria et al. 1997; Roseiro et al. 2003; Silva y Malcata 2005),
- ✓ el uso de leche de vaca concentrada mediante ultrafiltración (Low et al. 2006; Agboola et al. 2009);
- ✓ el incremento del tiempo de salado, Llorente y colaboradores (2014) concluyeron que los extractos de *C. scolymus* eran buenos sustitutos del cuajo animal para la elaboración industrial de queso tipo Gouda modificando el tiempo de salmuera;
- ✓ la inmovilización de las enzimas coagulantes (Esposito et al. 2016; Liburdi et al. 2018),
- ✓ la forma de preparación del extracto que influirá en: el tipo de enzimas presente (Vairo-Cavalli et al. 2013), en la microbiota (Gómez et al. 2001; Fernández-Salguero et al. 1999; Pouillet et al. 1991) y en la presencia de compuestos que pueden provocar reacciones no esperadas durante la elaboración del queso (Tamer 1993; Sousa y Malcata 1997; Silva et al. 2002; Fernández-Salguero et al. 2002; Chen et al. 2003b; Chazarra et al. 2007; Kazemipour et al. 2017; Ben Amira et al. 2017b); la dosis utilizada, la capacidad coagulante del material vegetal puede variar según el estadio en el que es recolectado, las regiones geográficas y la temporada (Cordeiro et al. 1994a; Barros et al. 2001; Galan et al. 2008) o el modo de almacenamiento (Tavaria et al. 2001).

En el caso de la especie *C. cardunculus* algunos de estos inconvenientes han sido solventados mediante un proceso de extracción de las proteasas desde fase acuosa y liofilización, que permite la obtención de un coagulante vegetal aséptico y de fácil estandarización y almacenamiento, el cual ha sido patentado (Fernández-Salguero et al. 2002). Hoy en día extractos del cardo estandarizados son comercializados bajo el nombre de Cynzime® por la empresa Fytozimus Biotech Inc., para la elaboración de quesos de

oveja o cabra, este producto también es distribuido por la empresa Enzyme Development Corporation (EDC) con el nombre Enzeco® Aspartic acid protease.

1.4.2 Proteasas aspárticas vegetales recombinantes y sus quesos

La clonación de las PAs en sistemas de expresión adecuados permite su caracterización y producción sin los inconvenientes asociados al uso de material vegetal (Tabla 8).

Tabla 8. Proteasas aspárticas vegetales recombinantes analizadas como coagulantes de la leche.

Especie vegetal	Enzima	Código GenBank	Hospedador	Referencias
<i>Cynara cardunculus</i>	Ciprosina B	X81984	<i>E. coli</i>	(White et al. 1999)
			<i>P. pastoris</i>	
			<i>S. cerevisiae</i>	(Sampaio et al. 2008)
			Suspensiones celulares de <i>C. cardunculus</i>	(Soares et al. 2000)
	VRen (cardosina B)	AJ237674	<i>K. lactis</i>	(Sampaio et al. 2010)
<i>Cirsium vulgare</i>	Procirsin	JN703462	<i>E. coli</i>	(Almeida et al. 2014; Almeida et al. 2017)
				(Lufrano et al. 2012)

Existen pocas PAs vegetales que hayan sido producidas mediante fermentación y evaluadas como coagulantes de la leche. Una de las proteasas que mayor atención ha centrado para su aplicación en la elaboración de quesos es la ciprosina B (Cordeiro et al. 1992), cuya producción ha sido investigada en varios hospedadores, entre ellos la cepa *E. coli* BL21(DE3)pLysS, utilizando el vector pET-3a, que daba lugar a la producción de la proteína recombinante en gran cantidad (10-20 mg/L de cultivo), pero en forma de agregados insolubles, a partir de los cuales los autores no encontraron un método eficaz para obtenerla correctamente plegada (White et al. 1999), sin embargo, cuando abordaron la expresión en la levadura *P. pastoris* utilizando el

vector pPICZ α , la cipsosina B se acumulaba en el medio de cultivo en su forma activa, con un rendimiento de 1mg de enzima por cada litro de cultivo. Esta enzima también ha sido clonada en la levadura *S. cerevisiae* (Soares 2000; Sampaio et al. 2008) y se ha investigado su producción a gran escala mediante cultivo en modo fed-batch (Sampaio et al. 2014) y mediante su clonación en otro sistema de expresión basado en la suspensión de cultivos celulares de *C. cardunculus* (Sampaio et al. 2010). La enzima obtenida a partir de los diferentes hospedadores presentaba unas características similares y actividad coagulante de la leche. La cipsosina B expresada en *S. cerevisiae* (Soares et al. 2000) produce quesos tipo manchego similares a los obtenidos con los extractos crudos del cardo, pero diferentes a los producidos con cuajo animal, sobre todo en lo relativo a la extensión de la proteólisis (Fernandez-Salguero et al. 2003). La empresa Fytozymus Biotech comercializa la enzima recombinante procedente del cardo bajo el nombre comercial Cyprozime™.

Otra proteasa aspártica del cardo, la cardosina B, sin la parte correspondiente al péptido señal, también ha sido expresada con fines industriales utilizando la levadura *K. lactis* (Almeida et al. 2014), que al igual que la levadura *S. cerevisiae*, utilizada para la producción de la cipsosina B, es un microorganismo GRAS (“Generally Regarded As Safe”) y por ello permitida para su uso en alimentación. El rendimiento de la producción fue mejorado (35x) mediante la eliminación de la región del gen correspondiente al PSI y la optimización del uso de codones. La enzima secretada al medio de cultivo tiene unas características proteolíticas similares a la cardosina B nativa. El medio de cultivo conteniendo la enzima recombinante optimizada, denominado por los autores VRen, demostró ser válido para coagular leche de oveja, cabra y vaca y su rendimiento quesero, en contra de lo que cabría esperar, es mayor que el obtenido con la FPC (Maxiren®).

Lufrano y colaboradores (2012) clonaron, expresaron y caracterizaron el ADNc que codifica para la PA cirsin, sin el péptido señal, en *E. coli* BL21 Star (DE3), utilizando el vector de expresión pET23-d. Al igual que ocurría con la cipsosina B (White et al. 1999), la proteasa se producía en forma de agregados insolubles, que mediante un proceso de plegado daban lugar a la obtención de la proteasa activa, sin necesidad de la eliminación del prosegmento, lo cual contradice lo observado para la mayoría de las PAs expresadas en un sistema de expresión procariótico.

1.5 Especies vegetales objeto de estudio

1.5.1 *Galium verum*

Descripción: La especie vegetal *Galium verum* fue descrita por Carlos Linneo en Species Plantarum, en el año 1753. Es una planta anual herbácea, de la familia *Rubiaceae*, que alcanza una altura de 50 a 120 cm, con tallo erecto y cuadrado. Las hojas estrechas y lanceoladas se agrupan en verticilos de 4-12 hojas. Las flores se disponen en inflorescencias de tipo panícula, de hasta casi 40 cm, con forma alargada. Las flores, de cáliz ausente, son de pequeño tamaño, con la corola de color amarillo vivo (Figura 7). La planta tiene un ligero olor a miel, que resulta más intenso en las flores que en otras partes. La floración se produce entre los meses de junio y agosto (Ortega-Olivencia y Alcaraz 2007).

Nombre: Su nombre científico hace referencia a su capacidad coagulante, en griego *gala*, significa leche, y *verum*, es un epíteto latino que significa "verdadero".

Entre los nombres vulgares por los que es conocida figuran:

- castellano: cuajaleches, presera, hierba sanjuanera, agana, hierba cuajadera
- gallego: agana, callaleite, herba do rodicio, preseira

- catalán: gali, herba de mató, herba colera, quallallet, espunyidella groga
- inglés: lady's bedstraw, maid's hair, chesse rennet, maid's hair, yellow cleavers.



Figura 7. *Galium verum*.

Distribución: Es una planta originaria de Europa y Asia. Crece hasta los 2600 m de altura, sobre suelos secos y pobres en nitrógeno, en bordes de caminos, prados, bosques poco espesos, matorrales, laderas, y esporádicamente cerca del mar. Se da en la mayor parte de Europa y de la Península Ibérica, aunque es menos frecuente en la zona noroeste.

Usos: *G. verum* es una planta sobre la que se han realizado pocos estudios, pero a la cual la medicina tradicional le atribuye una serie de propiedades terapéuticas, sin evidencias científicas de las mismas, como son: diurético, galactógeno, astringente en uso tópico, sedante digestivo, en cálculos y en arenillas renales, antiepiléptico. Antiguamente parece que fue utilizada para elaborar y dar color al queso de oveja y cabra (Grieve 1971) en determinadas zonas de

Inglaterra (Cheshire y Gloucestershire), teñir de rubio los cabellos y rellenar los colchones con sus flores, para que el olor de estas matase las pulgas.

A nivel científico se han identificado una serie de sustancias bioactivas en *G. verum*, como son: compuestos iridoides (Böjthe-Horváth et al. 1980; Zhao et al. 2009), flavonoides (Zhao et al. 2008), antraquinonas (Schulte et al. 1984; Zhao et al. 2006) y ácido clorogénico (Borisov et al. 1971). Se han estudiado sus propiedades antioxidantes (Mavi et al. 2004), antimicrobianas (Zhao et al. 2006) y anticancerígenas (Zhao et al. 2011; Schmidt et al. 2014).

1.5.2 *Ficus carica*

Descripción: La higuera es un arbusto o árbol pequeño, perteneciente a la familia *Moraceae*, que no suele superar los 6-8 metros de altura, con corteza grisácea y lisa; presenta una gran cantidad de células laticíferas que producen un látex lechoso, áspero y gomoso, que al entrar en contacto con el aire se espesa.

Las hojas son caducas, profundamente lobuladas, formadas por 3 o 5 lóbulos, miden de 10 a 20 cm de longitud y proactivamente igual de ancho, y se insertan en un peciolo largo y grueso.

Las flores se encuentran insertadas en el interior de un receptáculo carnoso, de forma piriforme, denominado sicono (Figura 8). Dentro del sicono están los frutos, denominados achenios. Florece a partir de la primavera y los higos maduran a finales de verano. La higuera originariamente monoica ha evolucionado a ginodioica (González-Rodríguez y Grajal-Martín 2011).

Nombre: Su nombre científico, *Ficus carica*, deriva de Caria, región de Asia Menor que destacaba por sus higos.

Distribución: El higo común es un árbol nativo del suroeste de Asia y del este del Mediterráneo. Hoy en día se cultiva en muchas regiones del mundo, como Turquía, Egipto, Marruecos, España,

Grecia, California, Italia, Brasil y otros lugares con inviernos típicamente suaves y veranos calurosos y secos (Chomé et al. 2011).

Usos: Los siconos se han usado tradicionalmente en alimentación y el látex parece que fue utilizado antiguamente para coagular la leche (Columela, De Re Rustica) (Faccia et al. 2012). Esta especie también se ha usado en medicina tradicional como laxante y para el tratamiento de enfermedades cardiovasculares y respiratorias por sus propiedades antiinflamatorias y antiespasmódicas.

Diversos estudios han revelado la presencia de numerosos compuestos bioactivos obtenidos desde diferentes partes de esta especie, compuestos como fitoesteroles, ácidos orgánicos, antocianinas, triterpenoides, compuestos fenólicos, cumarinas y compuestos volátiles como hidrocarburos, alcoholes alifáticos y otras clases de metabolitos secundarios. Se han realizado estudios científicos sobre su actividad antioxidante, anticancerígena, hepatoprotectora, antimicrobiana, hipoglucémica, hipolipidémica, antipirética, antimutagénica y otras (Mawa et al. 2013).



Figura 8. *Ficus carica*.



OBJETIVOS



2. OBJETIVOS

El objetivo que se persigue con el desarrollo de esta tesis es la caracterización de genes que codifican para proteasas aspárticas de las especies vegetales *Galium verum* y *Ficus carica* y su producción en sistemas heterólogos para ser aplicadas en la elaboración de quesos.

Para alcanzar este objetivo general se plantearon los siguientes objetivos específicos:

- 1) Comprobación de la actividad coagulante de la leche de extractos crudos de *G. verum* y *F. carica*.
- 2) Obtención de la secuencia génica de proteasas aspárticas de *G. verum* y *F. carica*.
- 3) Caracterización *in silico* de las proteasas codificadas por los genes secuenciados.
- 4) Expresión heteróloga de genes que codifican para proteasas aspárticas de *G. verum* y *F. carica*.
- 5) Caracterización de las proteasas aspárticas recombinantes de *G. verum* y *F. carica*.
- 6) Evaluación de las proteasas aspárticas obtenidas para su aplicación en la industria láctea





MATERIALES Y MÉTODOS



3. MATERIALES Y MÉTODOS

3.1. Materiales

3.1.1 Material vegetal

3.1.1.1 Recolección y almacenamiento

Plantas de la especie vegetal *G. verum* fueron recolectadas en la Sierra del Careón, localizada en la provincia de A Coruña, durante los meses de junio-julio, época de floración.

El material vegetal fue separado en dos lotes para su conservación: el primero fue desecado a temperatura ambiente y almacenado en un lugar seco y protegido de la luz solar; el segundo lote fue congelado en nitrógeno líquido y almacenado a -70 °C hasta su utilización.

El material vegetal procedente de *F. carica* fue recolectado en el ayuntamiento de Bueu, entre los meses de agosto-octubre. Se recogieron hojas, látex e infrutescencias ó siconos cuando estos estaban todavía inmaduros. La extracción del látex se realizó mediante una incisión longitudinal en los siconos y se introdujo en tubos de microcentrífuga. El material vegetal se congeló en nitrógeno líquido y se almacenó a -70 °C.

3.1.1.2 Obtención de extractos

Se prepararon extractos a partir de las diferentes partes (hojas, flores y tallos) de la planta seca *G. verum*. El material vegetal se mezcló con agua mineral (20%, p/v) y se mantuvo tres horas con agitación a temperatura ambiente. Transcurrido este tiempo la mezcla se trituró en un mortero, se centrifugó a 12000 g durante 20 min y se filtró a través de un papel de filtro Whatman #1 y posteriormente a través de un filtro de 0.22 µm.

El tejido vegetal fresco de *G. verum* (almacenado a -70 °C el mismo día que fue recolectado) se molió en un mortero en presencia

de nitrógeno líquido. El material pulverizado se mezcló con agua (20%, p/v), se centrifugó a velocidad máxima en una microcentrífuga y se filtró siguiendo el mismo procedimiento que el realizado con la planta seca.

En el caso de la especie *F. carica* el látex extraído a partir de higos verdes fue centrifugado a 5000 g durante 90 min a 4 °C para eliminar las gomas. El sobrenadante fue diluido 1/1000 con tampón fosfato sódico 10 mM pH 6.

Se midió el pH de todos los extractos obtenidos para comprobar que una vez añadidos a la solución de leche la coagulación no se produjese por acidificación.

3.1.2 Microorganismos

Las cepas bacterianas empleadas en este trabajo pertenecen a una única especie, *E. coli*. Las características genotípicas de las cepas utilizadas se detallan en la Tabla 9.

Tabla 9. Cepas de *Escherichia coli* utilizadas en este estudio.

Cepa	Genotipo
One Shot TOP 10	F- mcrA Δ(mrr-hsdRMS-mcrBC) φ80lacZΔM15 ΔlacX74 reA1 araD139 Δ(araleu) 7697 galU galK rpsL (StrR) endA1 nupG
DH5α	F- Φ80lacZΔM15 Δ(lacZYA-argF) U169 recA1 endA1 hsdR17 (rK-, mK+) phoA supE44 λ- thi-1 gyrA96 relA1
BL21(DE3)	F – ompT gal dcm hsdSB(rB- mB-) λ (DE3)
Rosetta-gami (DE3)	F – ompT hsdSB(rB- mB-) gal dcm lacY1 ahpC (DE3) gor522::Tn10 trxB pRARE (CamR, KanR, TetR)

- Las cepas **One Shot TOP 10** y **DH5α** presentan una elevada eficiencia de transformación y varias mutaciones que favorecen la estabilidad del ADN clonado. La delección

Φ80lacZΔM15 permite la α -complementación con el extremo N-terminal de la β -galactosidasa codificada por vectores del tipo pUC. Esta cepa fue utilizada para clonaciones rutinarias.

- La cepa **BL21 (DE3)** está diseñada para la expresión heteróloga de genes clonados en vectores que porten el promotor del fago T7. Presenta integrada en su genoma la secuencia que codifica para la T7 ARN polimerasa del fago λ DE3, cuya expresión es dependiente del promotor lacUV5 y del represor lacI; se induce por medio de la adición del inductor IPTG. Fue utilizada para la expresión del gen que codifica para la preprogalina A mediante el vector pET30a(+). Es defectiva en las proteasas OmpT y Lon, reduciendo la degradación de la proteína heteróloga.
- La cepa **Rosetta-gami (DE3)**, al igual que la cepa BL21(DE3) porta en su genoma el gen de la T7 RNA polimerasa. Esta cepa ha sido diseñada para facilitar la formación de puentes disulfuro en el citoplasma de *E. coli* mediante dos mutaciones (*trxB/gor*) y para aumentar la expresión de proteínas eucariotas mediante la expresión de seis ARNt para codones raros, que no se presentan naturalmente en *E. coli*. Fue utilizada para la expresión del gen que codifica para la preprogalina A y preprofigain A mediante el vector pET21a(+).

La cepa eucariota empleada fue *Pichia pastoris* **X-33**, levadura metilotrófica de genotipo silvestre (fenotipo Mut⁺: **M**ethanol **u**talization plus). Fue utilizada como hospedador para la expresión de los genes clonados utilizando los vectores pGAPZ α y pPICZ α .

La conservación de las cepas microbianas se realizó añadiendo 15-20 % de glicerol a un cultivo en fase exponencial y criopreservación a -70 °C.

3.1.3 Medios y condiciones de cultivo

3.1.3.1 *E. coli*

3.1.3.1 Medios de cultivo de *E. coli*

Los medios de cultivo empleados para *E. coli* fueron:

- **LB (Luria-Bertani)**

Triptona 1 %

Extracto de levadura 0.5 %

NaCl 1 %

El pH se ajustó a 7.5 con NaOH

Este medio fue suplementado con:

- Kanamicina (50µg/mL), cuando las células fueron transformadas con las construcciones realizadas con el vector pET-30a(+) o con el vector pCR®-Blunt II-TOPO®.

- Ampicilina (50 µg/mL) cuando las células fueron transformadas con el vector pUC19.

- IPTG (40 µL de una solución stock 20 mg/mL en dimetilformamida) y X-Gal (4 µL de una solución stock 200 mg/mL en agua destilada) según describe Sambrook y colaboradores (1989), para la detección de transformantes mediante el sistema de α -complementación en medio sólido cuando se utilizó el vector pUC19.

- Ampicilina (50 µg/mL), cloranfenicol (34 µg/mL), kanamicina (15 µg/ml), estreptomicina (50 µg/ mL) y tetraciclina (12.5 µg/mL), cuando las células fueron transformadas con las construcciones realizadas con el vector pET-21a(+).

- IPTG 1mM cuando fue necesaria la inducción de la expresión de los genes clonados en los vectores de la serie pET.

- Zeocina (25 µg/mL) cuando las células fueron transformadas con las construcciones realizadas con los vectores pGAPZ α A o pPICZ α A. Cuando el medio fue suplementado con zeocina la concentración

salina fue del 0.5 % y los cultivos se mantuvieron en oscuridad para evitar la pérdida de actividad de este antibiótico.

Los antibióticos fueron esterilizados mediante filtración empleando filtros de 0.22 μm (Millipore) antes de ser añadidos al medio estéril.

Los medios sólidos se elaboraron añadiendo agar al 2 %. El medio se esterilizó en autoclave con un ciclo de 15 min a 121 °C.

- **Medio SOC**

Tryptona	2 %
Extracto de levadura	0.5 %
NaCl	10 mM
KCl.....	2.5 mM
MgCl ₂	10 mM
MgSO ₄	10 mM
Glucosa	20 mM

Este medio de cultivo se empleó para realizar la transformación de la cepa One Shot TOP10 (Invitrogen), siguiendo las instrucciones del fabricante.

3.1.3.2 Condiciones de cultivo de *E. coli*

E. coli se cultivó de manera rutinaria a 37 °C y 180 rpm.

Para la producción de las proteasas recombinantes mediante las cepas de expresión BL21(DE3) y Rosetta-gami (DE3) se preparó un inóculo de 3 mL con una DO₆₀₀ próxima a 0.5, que se añadió a 100 mL de LB. El cultivo se mantuvo en agitador orbital a 180 rpm y 37 °C hasta que la DO₆₀₀ fue 0.5-1.0 (ca. 2-3 horas de cultivo). El cultivo de 100 mL se dividió en 2 matraces conteniendo 50 mL cada uno. Se añadió IPTG a uno de ellos, mientras el otro se utilizó como control

negativo de la expresión (la temperatura de cultivo y el tiempo de inducción se especifican en resultados).

3.1.3.2 *P. pastoris*

1.3.4 Medios de cultivo de *P. pastoris* X-33

- **YPD**

Extracto de levadura 1 %

Peptona 2 %

Dextrosa 2 %

El medio fue suplementado con ampicilina 100 µg/mL para evitar contaminaciones bacterianas; cuando los cultivos fueron destinados a la elaboración de queso no se añadió antibiótico.

- **YPD7**

Medio YPD tamponado a pH 7: se prepararon 900 mL de YPD conteniendo 10g de extracto de levadura, 20 g de peptona y 20 g de dextrosa y se esterilizaron mediante autoclave. Posteriormente se añadieron 100 mL de 1M tampón fosfato potásico pH 7 esterilizado por filtración (0.22 µm).

- **YPD4**

Medio YPD tamponado a pH 4: su preparación fue la misma que para el medio YPD7 salvo que el tampón añadido fue 1 M tampón citrato pH 4.

- **YPDS**

Medio YPD al que se le añadió sorbitol a una concentración final 1M.

- **BMGY/ BMMY** (Buffered Glycerol-complex Medium/ Buffered Methanol-complex Medium)
 - Extracto de levadura 1 %
 - Peptona 2 %
 - Tampón fosfato potásico, pH 6.0100 mM
 - YNB1.34 %
 - Biotina4 x 10⁻⁵ %
 - Glicerol (BMGY) 1 % ó 0.5 % methanol (BMMY)

- **BMG/ BMM** (Buffered Minimal Glycerol /Buffered Minimal Methanol)
 - Tampón fosfato potásico pH 6.0 100 mM
 - YNB 1.34 %
 - Biotina4 x 10⁻⁵ %
 - Glicerol (BMG) 1% o metanol (BMM) 0.5 %

- **MGY/MM** (Minimal Dextrose Medium/ Minimal Methanol)
 - YNB 1.34 %
 - Biotina4 x 10⁻⁵ %
 - Dextrosa (MD) 2 % o metanol (MM) 0.5 %

Los medios de cultivo (BMGY/BMMY, BMG/BMM y MD/MM) fueron preparados siguiendo las indicaciones del manual *Pichia* Expression kit (Invitrogen).

Los medios sólidos se elaboraron añadiendo agar al 1.5 %.

1.3.4 Condiciones de cultivo de *P. pastoris*

Las condiciones rutinarias de cultivo para *P. pastoris* fueron 28 °C y 230-300 rpm. La relación entre el volumen del medio de cultivo y la capacidad del matraz utilizado fue siempre inferior al 25 %.

Para la expresión de los genes bajo el control del promotor GAP (vector pGAPZ α) las cepas fueron cultivadas en YPD, YPD7 o YPD4 suplementado con ampicilina e inoculados a una DO₆₀₀ inicial de 0.02.

Para la expresión de los genes bajo el promotor AOX1 (vector pPICZ α) se inocularon 25 mL de medio MGY o BMG o BMGY con una colonia de *P. pastoris* conteniendo el inserto para la expresión del gen *gvIAP* y *fcIAP*. Esto se incubó hasta alcanzar una DO₆₀₀ de 2-6. Las células se centrifugaron a 3000 g durante 5 min a temperatura ambiente y se resuspendieron en medio MM, BMM o BMMY a una DO₆₀₀ de 1. Se añadió 0.5 % de metanol cada 24 horas para mantener la inducción.

Los cultivos se realizaron por triplicado.

3.1.4 Cebadores

Los cebadores empleados aparecen en la Tabla 10. Fueron diseñados empleando el programa informático Vector NTI Advance 10 (Invitrogen, Carlsbad, USA) y sintetizados por las empresas Sigma-Aldrich o IDT (Integrated DNA Technologies).

Tabla 10. Cebadores utilizados para PCRs, RACE PCRs y clonación.

Cebador	Aplicación	Secuencia (5'→3')
pAPF	Cebador degenerado	GAYACNGGNAGYTCYAAAYYTVTGG*
pAPR	Cebador degenerado	CCATMAANACRTCNCMMARRATCC*
cDNA Synthesis primer	Síntesis ADNc (RACE)	TTCTAGAAATTCAGCGGCCGC(T) ₃₀ VN
AP1	Adaptor primer 1 (RACE)	CCATCCTAATACGACTCACTATAGGGC
pMgv1f	gv1AP (3'RACE)	CAGGGGTTTCACCTGACAAAGGTCTGGCG
pMgv1r	gv1AP (5'RACE)	GCCATCTGGCAAAGCAGAACACATCGGTG
pMgv2f	gv2AP (3'RACE)	GTTCGTCTATTGCCGACTCTGGAAC TTC
pMgv2r	gv2AP (5'RACE)	TTGGCTCACAAACACCAGTAGCTCCA
pMfc1f	fc1AP (3'RACE)	AGGTTTTCATCGAGGGCTACACGAGAAGG
pMfc1r	fc1AP (5'RACE)	CTTC'TGCTGTCGGATCACGGTTAAGCC
pMfc2f	fc2AP (3'RACE)	GGGGAGCTCCAAATTTCTGGGTGCCATCG
pMfc2r	fc2AP (5'RACE)	GGCAACCAAAAACCTGGACACCAAGGCTCA
pGv1-EcoRI	Clonación en pET30a Clonación en pPICZα Clonación en pGAPZα	CCGGAAATTCGTGCTGAGAAATAGGTTTGAAAAAGAAAG
pGv1-XbaI	Clonación en pET30a Clonación en pPICZα Clonación en pGAPZα	GCTCTAGACTAAGCAGCCCTTCGCGAATCCTAG
F-NheIprofcl	Clonación en pET21a	CTAGCTAGCCTAAATGCGAATAAAACCTGAAGAAAGAGAG

(continuación Tabla 10)

R-NheIprofcl	Clonación en pET21a	CTAGCTAGCCTACGCAGCTAGAGCAAAACCCA
F-NheIprogv1	Clonación en pET21a	CTAGCTAGCGTGTGAGATAGGTTTGAAAAAGAAAGTC
R-NheIprogv1	Clonación en pET21a	CTAGCTAGCCTAAGCAGCCTTCGCGAATCC
pFc1-EcoRI	Clonación en pPICZα Clonación en pGAPZα	CCGGAAATTCCTAATGCGAATAAACCTGAAGAAGAGAGGCC
pFc1-XbaI	Clonación en pPICZα Clonación en pGAPZα	GCTCTAGACGCGCTCTACGCAGCTAGAGCAAAACCCAC
Fc1-EcoRI	Clonación en pPICZα Clonación en pGAPZα	CCGGAAATTCATGGAGGATATAGTACTTCTGAAGAACTATTTGGACGC
pGv2-EcoRI	Clonación en pGAPZαA	CCGAAATTCCTCCGACGCTTTGTTATCAGCATCTG
pGv2-XbaI	Clonación en pGAPZαA	GCTCTAGATCAAGCTGCTTCTGCGAATCCAAACCC
F17M13	Secuenciación	GTAAACGACGGCCAGT
R19M13	Secuenciación	GGAAACACAGCTATGACCATG
T7promotor	Secuenciación	TAATACGACTCACTATAGG
T7terminador	Secuenciación	TATGCTAGTTATTGCTCAG
5'AOX1	Secuenciación	GACTGGTTCCAATTGACAAGC
3'AOX1	Secuenciación	GCAATGGCAATCTGACATCC
pGAP	Secuenciación	GTCCCTATTTCGAATCAATTGAA

*Y= C, T; R= A, G; V= A, G, C; M= A, C; N= A, C, G, T

Las dianas de las enzimas de restricción empleadas para la clonación aparecen subrayadas.

3.1.5 Vectores

3.1.5.1 pCR[®]-Blunt II-TOPO[®]

El vector pCR[®]-Blunt II-TOPO[®] (Figura 9) es suministrado por el proveedor en forma lineal, con la ADN topoisomerasa I del virus *Vaccinia* unida covalentemente a los extremos 3' de cada hebra de ADN. Permite la clonación de productos de PCR con extremos romos sin necesidad de utilizar ligasa y la selección directa de recombinantes por interrupción de la expresión del gen letal de *E. coli ccdB*, gen que se encuentra unido por el extremo C-terminal a un fragmento del gen *lacZα*. La unión del inserto interrumpe la expresión del producto de fusión *lacZα-ccdB*, permitiendo que solo crezcan las colonias recombinantes. Además, este vector permite seleccionar transformantes resistentes a kanamicina [kan(R)] y a zeocina [Zeo(R)]. El inserto se une a una región que está flanqueada por dos dianas de restricción para la enzima *EcoRI*, lo cual facilita la posterior liberación del inserto. Las secuencias F17M13 y R19M13 (Tabla 10) permiten la secuenciación del inserto.

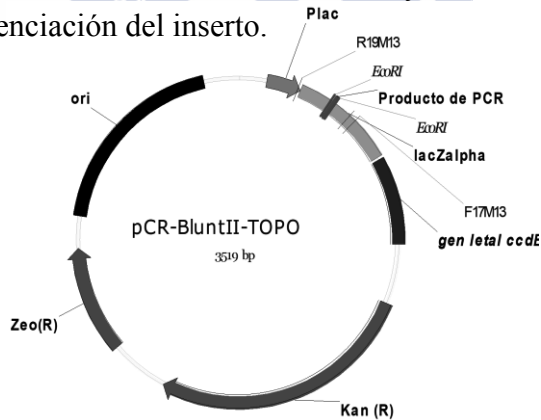


Figura 9. Dibujo esquemático del vector pCR[®]-Blunt II-TOPO[®]. Los elementos que lo componen son: origen de replicación en *E. coli* (ori); promotor/operador del operón lactosa (Plac); fragmento del gen *lacZα* fusionado al gen letal *ccdB*; gen de resistencia a la kanamicina [kan(R)]; gen de resistencia a la zeocina [Zeo(R)]. La unión del producto de PCR interrumpe la pauta de lectura del gen *ccdB*. Los cebadores R19M13 y F17M13 fueron utilizados para la secuenciación.

3.1.5.2 pUC19

El vector pUC19 (Invitrogen) (Figura 10) contiene el gen de resistencia a la ampicilina [Amp(R)], que permite la selección de los transformantes. Los insertos clonados en el sitio de clonaje múltiple (MCS) pueden ser expresados bajo el control del promotor *lac* (Plac), inducible por IPTG. Además, contiene el gen *lacZ*, que permite seleccionar por color las colonias recombinantes en medio de cultivo con X-Gal e IPTG (Yanisch-Perron et al. 1985). Contiene el origen de replicación de *E. coli*.

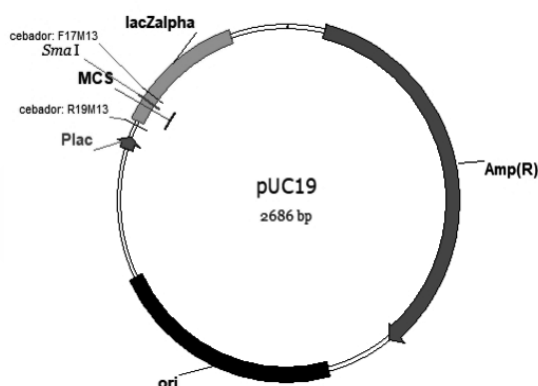


Figura 10. Dibujo esquemático del vector pUC19. Los elementos que lo componen son: origen de replicación (ori); promotor/operador del operón lactosa (Plac); un sitio de clonaje múltiple (MCS), fragmento del gen lacZalpha, que codifica para la enzima β -galactosidasa, y gen de resistencia a la ampicilina [Amp(R)]. Los cebadores R19M13 y F17M13 fueron utilizados para la secuenciación.

3.1.5.3 pET

En el sistema de expresión pET los genes *gvIAP* o *fcIAP* fueron clonados bajo el control del promotor del bacteriófago T7. La acción del promotor T7 depende de la ARN polimerasa del fago T7, cuyo gen (T7 gen 1) se encuentra insertado en el genoma de las células de *E. coli* BL21(DE3) o Rosetta-gami (DE3), bajo el control del promotor *lac*, inducible por IPTG (Figura 11).

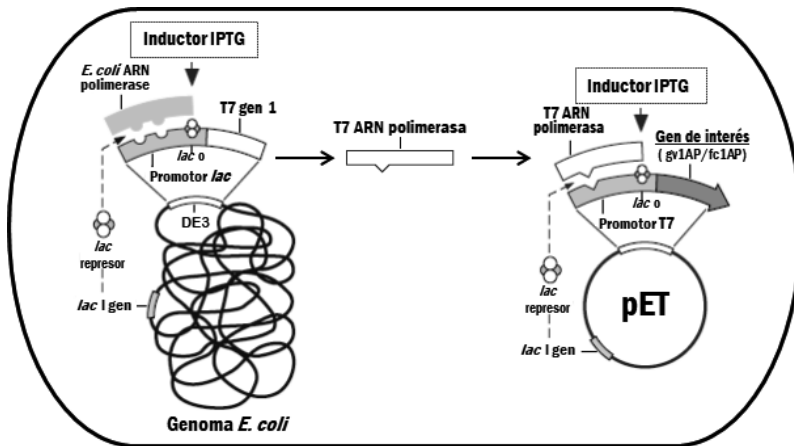


Figura 11. Esquema del sistema de expresión pET (Novagene).

3.1.5.3.1 pET30a(+)

El vector de expresión pET30a(+) (Figura 12) contiene el promotor T7.

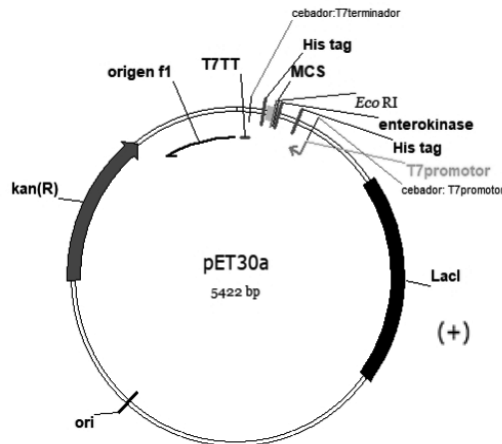


Figura 12. Dibujo esquemático del vector pET30a(+). Los elementos que lo componen son: origen de replicación (**ori**); gen de resistencia a la kanamicina [**kan(R)**]; el origen de replicación del bacteriófago F1 (**origen f1**); el terminador de la transcripción T7 (**T7TT**); un sitio de clonaje múltiple (**MCS**); cola de histidinas (**His tag**) que permite purificar y detectar la proteína recombinante; sitio de escisión para la enterocinasa; promotor del bacteriófago T7; el gen represor del operón lac (**LacI**). Los cebadores **T7terminador** y **T7promotor** fueron utilizados para la secuenciación.

Este plásmido incluye el operador lac y el gen lacI, para utilizar IPTG como inductor de la expresión, y el gen de resistencia a la kanamicina [kan(R)] como marcador de transformación. Contiene una cola de histidinas (His tag), que facilita la purificación de la proteína recombinante mediante cromatografía de afinidad sobre metal inmovilizado (IMAC).

3.1.5.3.2 pET21a(+)

El vector de expresión pET21a(+) (Figura 13) contiene el promotor T7. Este plásmido incluye el operador lac y el gen lacI para utilizar IPTG como inductor de la expresión y el gen de resistencia a la ampicilina como marcador de la transformación [Amp(R)].

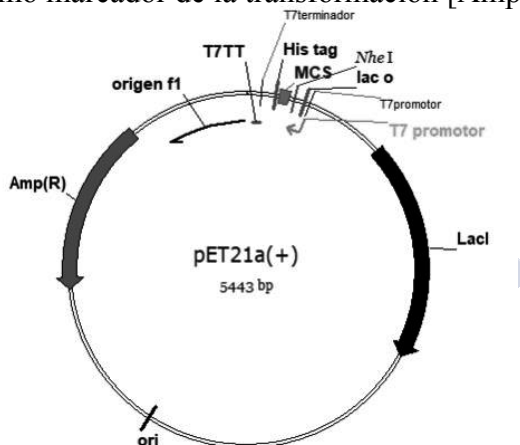


Figura 13. Dibujo esquemático del vector pET21a(+). Los elementos que lo componen son: origen de replicación (ori); gen de resistencia a la ampicilina [Amp(R)]; el origen de replicación del bacteriófago F1 (origen f1); el terminador de la transcripción T7 (T7TT); cola de histidinas (His tag) que permite purificar y detectar la proteína recombinante; un sitio de clonaje múltiple (MCS); operador lac (lac o); sitio de escisión para la enterocinasa; promotor del bacteriófago T7; el gen represor del operón lac (LacI). Los cebadores T7terminador y T7promotor fueron utilizados para la secuenciación.

3.1.5.4 pPICZα A

pPICZα A (Invitrogen, Carlsband, USA) (Figura 14) es un vector integrativo que posee el promotor del gen alcohol oxidasa (AOX1), inducible cuando la levadura crece en presencia de metanol como

única fuente de carbono. Este vector produce la proteína recombinante fusionada en su extremo N-terminal a la señal de secreción factor- α de *S. cerevisiae*, lo cual permite obtener la proteasa recombinante en el sobrenadante del medio de cultivo. La selección de los transformantes está basada en la zeocina, que es bifuncional en *P. pastoris* y *E. coli*.

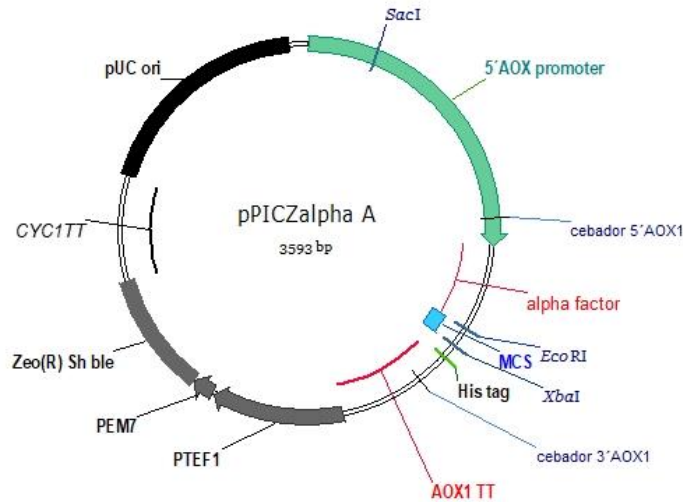


Figura 14. Dibujo esquemático del vector pPICZ α A. Este vector presenta los siguientes elementos: promotor del gen de la alcohol oxidasa (5' AOX promoter) inducible por metanol; secuencia señal factor- α (alpha-factor) para la expresión extracelular de la proteína recombinante; un sitio de clonaje múltiple (MCS); una cola de histidinas (His tag) que permite purificar y detectar la proteína recombinante; el terminador de la transcripción del gen AOX1 (AOX1TT); el gen de *Streptoalloteichus hisdustanus* que confiere resistencia a la zeocina (gen *Sh ble*) bajo dos promotores, uno procariota (PEM7), que dirige la expresión del gen en *E. coli* y otro eucariota (PTEF1), que dirige la expresión en *P. pastoris*; la región de terminación de la transcripción del gen *CYC1* de *S. cerevisiae*, que permite estabilizar y procesar el ARNm del gen *Sh ble*; el sitio de replicación del plásmido en *E. coli* (pUC ori). Los sitios de restricción EcoRI y XbaI pertenecientes al MCS fueron utilizados para insertar los genes que codifican para las proteasas aspárticas vegetales; el sitio SacI localizado en el locus AOX1 fue utilizado para linearizar el plásmido y conseguir la integración de la construcción en el genoma de *P. pastoris*. Los cebadores 5' AOX1 y 3' AOX1 fueron utilizados para la secuenciación.

3.1.5.5 pGAPZα A

El vector pGAPZα A (Invitrogen, Carlsband, USA) (Figura 15) es un vector integrativo, que ha sido diseñado para la expresión constitutiva de proteínas en *P. pastoris* bajo el promotor del gen de la gliceraldehído-3-fosfato deshidrogenasa (pGAP). Este vector produce la proteína recombinante fusionada en su extremo N-terminal a la señal de secreción factor-α de *S. cerevisiae*, lo cual permite obtener la proteasa recombinante en el sobrenadante del medio de cultivo. La selección de los transformantes está basada en la zeocina, que es bifuncional en *P. pastoris* y *E. coli*.

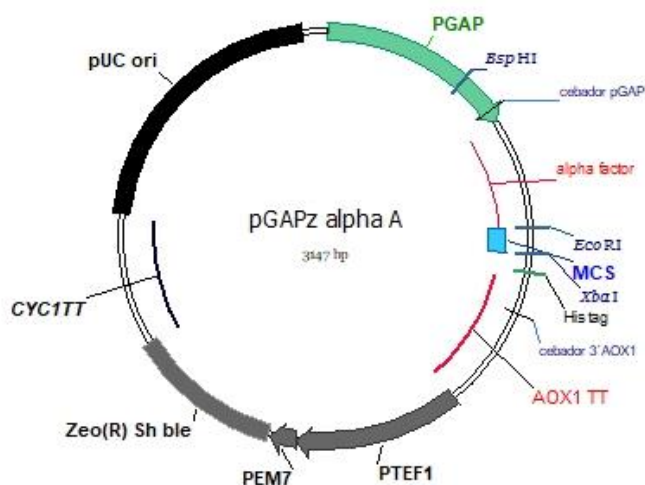


Figura 15. Dibujo esquemático del vector pGAPZα A. Los elementos que lo componen son los mismos que los presentes en el vector pPICZα A (Figura 14) a excepción del promotor utilizado para la expresión, promotor del gen gliceraldehído-3-fosfato deshidrogenasa (PGAP); el sitio BspHI fue utilizado para linearizar el plásmido y conseguir la integración de la construcción en el genoma de *P. pastoris*. Los cebadores pGAP y 3' AOX1 fueron utilizados para la secuenciación.

3.2 Técnicas de manipulación de ácidos nucleicos

3.2.1 Extracción y purificación de ácidos nucleicos

3.2.1.1 Extracción de ARN desde tejidos vegetales

La extracción se realizó con el kit comercial RNeasy® Plant Mini Kit (Qiagen, Hilden, Germany) siguiendo las instrucciones del fabricante.

Para evitar la degradación del ARN durante el proceso se llevó a cabo un tratamiento previo del material utilizado para la extracción con anti-RNasas:

- El agua se trató con DEPC (dietilpirocarbonato): se disolvieron 0.1 mL de DEPC en 100 mL de agua. La solución se mantuvo a 37 °C durante 12 horas, transcurridas las cuales se esterilizó en autoclave a 121 °C durante 30 min.
- El material plástico y de cerámica, así como las superficies de trabajo se trataron con NaOH 0.1 M y EDTA 1 mM disuelto en agua previamente tratada con DEPC.
- El material de cristal se sumergió 12 h a 37 °C en agua con 0.1 % de DEPC, tras lo cual se esterilizó en autoclave a 121 °C durante 30 min.
- Los tanques de electroforesis se lavaron con SDS al 0.5 % disuelto en agua tratada con DEPC. A continuación, se lavaron con agua sin el detergente y por último se lavaron con etanol.

Los pasos seguidos para la extracción fueron:

- El material vegetal almacenado a -70 °C fue molido en un mortero de cerámica, empleando nitrógeno líquido. Una vez reducido a polvo, la muestra se introdujo en un tubo preenfriado de 2 mL al que se añadieron 450 µL de tampón de lisis RLT conteniendo 10 µL/mL de β-ME (RNeasy® Plant Mini Kit, Qiagen) y se incubó durante 1-3 min a 56 °C para favorecer la lisis. El lisado se transfirió a una columna QIAshredder colocada en un tubo colector de 2 mL y centrifugó 2 min a 18000 g para eliminar los restos celulares y homogenizar el lisado.

La columna fue desechada y se transfirió el sobrenadante desde el tubo colector a un nuevo tubo.

- Se añadieron 0.5 volúmenes de etanol 100 % al lisado y se mezcló con una pipeta. Esta mezcla se introdujo en una mini columna (RNeasy spin column, Qiagen) colocada en un tubo colector y se centrifugó durante 15 s a 12000 g.

- Se desechó el contenido del tubo colector y la columna conteniendo el ARN en la matriz se emplazó en el mismo tubo y se lavó con 700 μ L de tampón RW1 (RNeasy Plant Mini Kit, Qiagen). Se realizó una centrifugación durante 15 s a 10000 g y se desechó el tubo colector y su contenido. Se añadieron 500 μ L de tampón RPE (RNeasy Plant Mini Kit, Qiagen) a la columna emplazada en un nuevo tubo y se centrifugó en las mismas condiciones. Se realizó un último lavado con este tampón y centrifugó durante 2 min a 12000 g.

- La elución del ARN retenido en la columna se realizó emplazando la columna en un nuevo tubo, añadiéndole 30 μ L de agua libre de RNasas (Invitrogen) y realizando una centrifugación durante 1 min a 12000 g. El proceso se repitió utilizando el eluido como eluyente para obtener el ARN más concentrado.

Se eliminó cualquier traza de ADN con DNasa I (1 U por cada μ g de ARN) (DNase I, Amp Grade; Invitrogen) utilizando el tampón de reacción Tris-HCl 20 mM, pH 8.4, MgCl₂ 2 mM, KCl 500 mM e incubación durante 15 min a temperatura ambiente. La DNasa se inactivó mediante la adición de 1 μ L de una solución EDTA 25 mM por cada 10 μ L de mezcla de reacción e incubación durante 10 min a 65 °C.

Cuando fue necesario las muestras de ARN se almacenaron a -20 °C hasta su uso.

3.2.1.2 Extracción de ARN desde *P. pastoris*

La extracción de ARN se realizó utilizando el kit RNeasy Plant Mini Kit (Qiagen), siguiendo las instrucciones del fabricante. Se utilizó DNase I para eliminar cualquier traza de ADN que pudiese contener el ARN obtenido.

3.2.1.3 Extracción de ADN genómico desde *P. pastoris*

La extracción de ADN genómico desde *Pichia* se realizó utilizando el Kit MasterPure™ Yeast DNA Purification kit (Epicentre® Biotechnologies, Madison USA) siguiendo las instrucciones del fabricante.

3.2.1.4 Extracción de ADN genómico desde *G. verum*

La extracción del ADN genómico se realizó a partir de material vegetal, principalmente flores, pulverizado en un mortero utilizando nitrógeno líquido. Se mezclaron en un tubo de microcentrífuga 0.2 g del pulverizado con 500 µL de tampón de extracción [Tris HCl 100 mM pH8, CTAB 2 %, PVPP 2 % (Sigma, USA), NaCl 1.4 M, EDTA 20 mM, β-ME 0.6 mM], la mezcla se incubó a 65 °C durante 2 h con agitación. Transcurrido este tiempo la muestra se centrifugó (14000 g, 8 min) para eliminar los restos celulares y se agregaron 400 µL de cloroformo: alcohol isoamílico (24:1). Transcurridos 5 min a temperatura ambiente la mezcla se centrifugó (1500 g, 10 min) y la fase superior se transfirió a un nuevo tubo e incubó con 3 µL de RNase A (Sigma) (10 mg/mL) a temperatura ambiente durante 30 min, transcurridos los cuales se añadió un volumen de fenol: cloroformo: alcohol isoamílico (25:24:1), saturado con Tris 10 mM, pH 8.0, EDTA 1mM (Sigma) y se realizó una centrifugación a 1500 g durante 10 min. Se recogió la fase superior y se repitió la extracción. Finalmente, para precipitar el ADN, se añadieron 0.6 volúmenes de isopropanol y se realizó una centrifugación a 1500 g durante 10 min. Se eliminó el sobrenadante y se realizó un lavado con etanol al 70 %.

El ADN precipitado se resuspendió en H₂O_{mQ} y se determinó su concentración por espectrofotometría.

3.2.1.5 Extracción de ADN plasmídico desde *E. coli*

El aislamiento de pequeñas cantidades de ADN plasmídico (1-5 µg) se realizó a partir de cultivos crecidos en 4 mL de LB, suplementado con el antibiótico de selección adecuado, utilizando el kit “Miniprep Express matrix” (MP Biomedicals, Ohio, USA) o el kit “High Pure Isolation Kit” (Roche Applied Science, Indianapolis, USA).

El aislamiento de mayores cantidades de ADN plasmídico (> 100 µg) se realizó a partir de cultivos crecidos en 50-100 mL de LB, suplementado con el antibiótico de selección adecuado, utilizando el kit “Plasmid Midiprep” (Qiagen, Hilden, Germany).

3.2.1.6 Purificación de ADN desde geles de agarosa

Los fragmentos de ADN separados mediante electroforesis en geles de agarosa se purificaron cortando las bandas de interés con una cuchilla estéril desde el gel mientras estas se visualizaban con luz ultravioleta. El ADN se extrajo empleando el kit Wizard PCR preps Purification System (Promega) o el kit “Illustra GFX PCR DNA and Gel band Purification kit” (GE Healthcare), siguiendo las instrucciones del fabricante.

3.2.1.7 Purificación de ADN desde reacción enzimática

La purificación del ADN desde PCR o reacciones de digestión enzimática se realizó utilizando el kit “Illustra GFX PCR DNA and Gel Band Purification Kit” (GE Healthcare) siguiendo las instrucciones suministradas por el fabricante.

3.2.2 Digestión enzimática del ADN

La digestión del ADN se realizó mediante el empleo de enzimas de restricción comerciales (Takara Bio. Inc. Japan o New

England BioLabs, Beverly, MA) siguiendo las recomendaciones del fabricante. Cuando fue necesario realizar una digestión doble, con enzimas cuyos tampones no eran compatibles, se utilizó un tampón universal o se realizó la purificación del ADN previa al corte con la segunda enzima.

La inactivación de las enzimas se realizó siguiendo las instrucciones del fabricante.

Cuando fue necesario se consultó la base de datos sobre enzimas de restricción REBASE (<http://rebase.neb.com/rebase/rebase.html>)

3.2.3 Defosforilación del ADN

Cuando fue necesario el ADN se defosforiló con la fosfatasa alcalina de camarón (*Shrimp Alkaline Phosphatase*: SAP) (Takara o Promega), siguiendo el protocolo suministrado por el fabricante.

3.2.4 Fosforilación del ADN

Cuando fue necesario el ADN se fosforiló con la enzima T4 polinucleótido quinasa (Takara), siguiendo el protocolo suministrado por el fabricante.

3.2.5 Ligación del ADN

La ligación entre los plásmidos linearizados y fragmentos de ADN, obtenidos por digestión enzimática o mediante PCR, se realizó en relaciones molares 1:1, 1:2 y 3:3 utilizando la ADN polimerasa T4 (Promega, Madison, WI).

Cuando los extremos eran romos la mezcla de ligación se incubó a 16 °C durante toda la noche, mientras que cuando los extremos eran cohesivos se mantuvo a 22 °C durante 3 horas.

La ligación de entre fragmentos de extremos romos y el vector pCR®-Blunt II-TOPO® se realizó mezclando ambos en una solución salina e incubando la mezcla 5 min a 22-23 °C.

3.2.6 Secuenciación

La secuenciaciones de ácidos nucleicos fueron realizadas por la empresa Sistemas genómicos (https://www.sistemasgenomicos.com/web_sg/index.php) o la empresa Secugen (<https://www.secugen.es/>), utilizando cebadores específicos enviados con la muestra (Tabla 10) o proporcionados por la empresa.

3.2.7 Síntesis de ácidos nucleicos

3.2.7.1 Síntesis de ADN complementario

El ADNc se sintetizó utilizando la transcriptasa inversa del virus de la leucemia murina de Moloney (BioScript MMLV; Bioline, London, UK) o bien la transcriptasa inversa AMV del virus de la mieloblastosis aviar suministrada en el kit 1st Strand cDNA Synthesis Kit for RT-PCR (Roche).

Los cebadores empleados fueron: Oligo-p(dT)₁₅, Random p(dN)₆ (ambos suministrados con el kit) o pAPR (específico, Tabla 10).

Los componentes de la mezcla de reacción cuando se utilizó la AMV transcriptasa inversa (Roche) fueron: tampón de reacción (Tris 10 mM, KCl 50 mM, pH 8.3), MgCl₂ 5 mM, dNTPs 1 mM, cebador [Oligo-p(dT)₁₅ (1.6 µg) o Random p(dN)₆ (3.2 µg) o pAPR (10 pM)], inhibidor de RNasa (50 U), transcriptasa inversa AMV (20 U), ARN* (≤1 µg), en un volumen final de 20 µL.

Cuando se empleó la transcriptasa inversa MMLV (BioScript) los componentes de la mezcla de reacción fueron: tampón de reacción 1X (suministrado con la enzima), dNTPs 0.5 mM, cebador [Oligo-p(dT)₁₅ (0.8 µg) o Random p(dN)₆ (1.6 µg) o pAPR (1 pM)], transcriptasa inversa MMLV (200 U), ARN* (≤5 µg) en un volumen final de 20 µL.

* Las muestras de ARN fueron desnaturalizadas a 65 °C durante 15 min y emplazadas en hielo antes de iniciar la reacción con el fin de eliminar las estructuras secundarias.

En algunas muestras no se añadió transcriptasa inversa para ser utilizadas posteriormente como control negativo.

El programa empleado en el termociclador para la síntesis del ADNc fue:

1. 25 °C, 10 min (hibridación cebadores)
2. 42 °C, 60 min (transcripción inversa)
3. 99 °C, 5 min (desnaturalización AMV-RT), o 70 °C, 10 min (desnaturalización MMLV-RT)
4. 4 °C, 5 min (enfriamiento)

El ADNc fue utilizado en el momento o almacenado a -20 °C hasta su uso.

3.2.7.2 Síntesis de ADN mediante PCR

El ADN se amplificó mediante reacción en cadena de la polimerasa (PCR). Todas las reacciones se llevaron a cabo en un termociclador MyCycler™ Thermal Cycler System (BioRad). Las polimerasas utilizadas fueron:

- ✓ **Biotaq** (Bioline, London, UK) para llevar a cabo las amplificaciones rutinarias.
- ✓ **Accuzyme** (Bioline, London UK), **PfuUltra®** (Stratagene) o **PrimeSTAR® HS DNA Polymerase** (Takara) cuando se necesitó una alta fidelidad de copia. Estas enzimas se utilizaron siempre que el producto de la PCR estaba destinado a la clonación en el vector pCR®-Blunt II-TOPO® (Invitrogen).

Los componentes y las condiciones de cada reacción fueron los recomendados para cada enzima variando la temperatura de hibridación en función de los cebadores utilizados.

La mezcla de reacción cuando la amplificación se realizó con la enzima **Biotaq** consistió en: tampón NH₄ (suministrado con la

enzima), MgCl_2 2 mM, dNTP 0.25 mM (cada uno), 2.5 U Biotaq, ADN (1-500 ng), cebadores 0.2 μM (cada uno) y $\text{H}_2\text{O}_{\text{mQ}}$ (csp 50 μL). Cuando los cebadores utilizados fueron pAPF/pAPR se añadió BSA 0.001 % .

La mezcla de reacción cuando la amplificación se realizó con la enzima **Accuzyme** consistió en: tampón Accubuffer (suministrado con la enzima), MgCl_2 0.5-5 mM, dNTP 0.25 mM (cada uno), 2.5-3.75 U Accuzyme, cebadores 0.2 μM (cada uno); ADN (1-500 ng) y $\text{H}_2\text{O}_{\text{mQ}}$ (csp 50 μL).

La mezcla de reacción cuando la amplificación se realizó con la enzima **PfuUltra** consistió en: tampón suministrado con la enzima, dNTP 0.25 mM, cebadores 0.3 μM (cada uno), 2.5 U PfuUltra, ClMg^{+2} 1 mM, ADN (1-500 ng).

La mezcla de reacción cuando la amplificación se realizó con la enzima **PrimeSTAR® HS** consistió en: tampón suministrado con la enzima, dNTP 0.2 mM, cebadores 0.3 μM , 1.5 U PrimeSTAR® HS, ADN (< 200ng).

La temperatura de hibridación y los ciclos utilizados en cada reacción se especifican en los apartados correspondientes con la siguiente notación: (temperatura de desnaturalización inicial tiempo)_{1x}; (temperatura de desnaturalización tiempo, temperatura de hibridación tiempo, temperatura de extensión tiempo)_{n°ciclos}; (temperatura extensión final tiempo)_{1x}

3.2.7.3 RT-PCR de las cepas de *P. pastoris* recombinantes

El ARN se extrajo utilizando el RNeasy® Plant Mini Kit (Qiagen, Hilden, Germany) a partir de cultivos de *P. pastoris* transformados con las construcciones realizadas con el vector pGAPZ α cultivadas durante 24 horas en las condiciones habituales, o con las realizadas con el vector pPICZ α transcurridos cuatro días de inducción con metanol. Para eliminar el ADN contaminante se realizó un tratamiento con DNasa I (Invitrogen).

La síntesis de la hebra complementaria se realizó utilizando el Transcriptor High Fidelity cDNA Synthesis Kit (Roche); se prepararon muestras a las que no se le añadió transcriptasa inversa para ser usadas como controles.

La PCR se realizó utilizando los cebadores (Tabla 10): pMgv1f/pMgv1r en el caso de la cepa transformada con la construcción pPICZ α -progv1AP y pMfc1f/ pMfc1r en el caso de las cepas transformadas con las construcciones pPICZ α -profc1AP y pPICZ α -profc1AP. La temperatura de hibridación empleada fue de 68 °C.

En el caso de las construcciones pGAPZ α los cebadores utilizados fueron (Tabla 10): pGv1-EcoRI/pGv1-XbaI para la cepa transformada con la construcción pGAPZ α -progv1AP y pFc1-EcoRI/pFc1-XbaI para la cepa transformada con la construcción pGAPZ α -profc1AP. La temperatura de hibridación fue de 55 °C.

3.2.8 Electroforesis

3.2.8.1 Electroforesis de ARN

Con objeto de comprobar la integridad y calidad de las muestras de ARN obtenidas desde las especies vegetales se llevó a cabo una electroforesis en gel desnaturalizante (agarosa-formaldehído).

La composición del gel fue agarosa 1.2 % (p/v) en tampón FA [3-(N-Morpholino) propanesulfonic acid (MOPS) 20 mM, acetato sódico 5 mM, EDTA 1 mM] con el pH ajustado a 7.0 mediante adición de NaOH. La mezcla fue licuada utilizando un microondas y se introdujo en un baño a 65 °C, cuando alcanzó esa temperatura se añadió formaldehído hasta una concentración final 0.25 M y bromuro de etidio (0.1 µg/mL). El desarrollo de la electroforesis tuvo lugar en tampón FA suplementado con formaldehído 0.25 M, con una diferencia de potencial de 6 V/cm. Para preparar la muestra se añadió un volumen de tampón de carga [azul de bromofenol 0.16 %, MOPS 80 mM, acetato sódico 20 mM, EDTA 4 mM, formaldehído 0.88 M,

glicerol 20%, formamida 30 % (v/v) y FA 40% 10X] por cada cuatro volúmenes de muestra y se mantuvo la mezcla a 65 °C entre 3-5 min para desnaturalizarla, transcurridos los cuales se enfrió en hielo y cargó en el gel. El ARN se visualizó en un transiluminador de luz UV.

3.2.8.2 Electroforesis de ADN

Las moléculas de ADN se separaron mediante electroforesis en geles de agarosa. La concentración de agarosa utilizada varió entre 0.9-2 % (p/v), dependiendo del tamaño del ADN a separar. La preparación de los geles y el desarrollo de las electroforesis se realizaron en tampón TAE (Tris acetato 40 mM, EDTA 1 mM, pH 8). En los geles se incorporó bromuro de etidio para la visualización del ADN mediante luz ultravioleta. Las muestras se mezclaron con una solución de tampón de carga comercial (Promega). La electroforesis se llevó a cabo en una cubeta horizontal a un voltaje entre 1-5 V/cm. La estimación del tamaño de las bandas de ADN se realizó utilizando marcadores de pesos moleculares comerciales como 2-Log DNA Ladder (New England Biolabs) o 1 kb ladder (Promega).

El ADN se visualizó iluminándolo con luz ultravioleta en un transiluminador con Sistema de documentación de geles ChemiDoc XRS+ (BioRad) y se analizó utilizando el software Bio-Rad Image Lab™.

3.2.9 Determinación de la concentración de ácidos nucleicos

3.2.9.1 Cuantificación mediante espectrofotometría

La cuantificación de ADN y ARN mediante espectrofotometría se realizó empleando una cubeta de cuarzo de 20 µL de capacidad y un espectrofotómetro Biophotometer (Eppendorf). Como blanco se utilizó H₂O_{mQ}. Se consideró que una unidad de absorbancia a 260 nm equivale a 44 µg de ARN o 50 µg de ADN por cada mL de muestra.

Cuando el volumen de ácidos nucleicos que se quería destinar a la cuantificación era pequeño (1-2 µL) se utilizó el espectrofotómetro

NanoDrop 1000 Spectrophotometer (Thermo Scientific, USA) siguiendo las instrucciones del manual.

Las relaciones A_{260}/A_{280} y A_{260}/A_{230} se tuvieron en cuenta para valorar la pureza de los ácidos nucleicos.

3.2.9.2 Cuantificación mediante electroforesis

La cuantificación de ácidos nucleicos desde gel se realizó comparando la intensidad de la banda de ADN a determinar con la de las bandas de concentración conocida del marcador de pesos moleculares, utilizando un sistema de documentación de geles ChemiDoc XRS+ (BioRad) y el software Bio-Rad Image Lab™.

3.2.10 Amplificación de genes que codifican para PAs de *G. verum* y *F. carica*

3.2.10.1 PCR mediante cebadores degenerados

La amplificación de secuencias internas de genes que codifican para PAs de *G. verum* y *F. carica* se realizó mediante PCR utilizando los cebadores degenerados (pAPF y pAPR, Tabla 10), diseñados en base a secuencias conservadas de PAs vegetales (Laloi et al. 2002).

En primer lugar, se realizó una PCR, con gradiente de temperatura, utilizando la ADN polimerasa Biotaq (Bioline) y como molde el ADNc sintetizado a partir de ARN total extraído desde *G. verum* o de *F. carica* (apartado 3.2.7.1).

EL programa de amplificación empleado fue:

✓ (94 °C 4 min)_{1x} (94 °C 30 s, 40-55 °C 30 s, 70 °C 2 min)_{30x} (70 °C, 7 min)_{1x}

El producto de estas amplificaciones fue analizado mediante electroforesis en gel de agarosa al 1 % y las bandas de tamaño ~1Kb purificadas y reamplificadas mediante una PCR realizada con la ADN polimerasa Accuzyme (Bioline), que a diferencia de la anterior deja extremos romos, necesarios para la introducción del producto de PCR en el vector pCR®-Blunt II-TOPO® (Invitrogen, Carlsbad, CA, USA).

Los programas de amplificación empleados fueron:

- ✓ *G. verum*: (94 °C 4 min)_{1x} (94 °C 30 s, 46 °C 30 s, 70 °C 2 min)_{30x} (70 °C 7 min)_{1x}
- ✓ *F. carica*: (94 °C 4 min)_{1x} (94 °C 30 s, 43 °C 30 s, 68 °C 1.30 min)_{30x} (68 °C 7 min)_{1x}

Los fragmentos amplificados fueron purificados desde el gel y clonados en el vector pCR®-Blunt II-TOPO®. El ADN plasmídico fue extraído desde varios clones y posteriormente secuenciados utilizando los cebadores F17M13 y R19M13.

En base a las secuencias obtenidas se diseñaron cebadores específicos para obtener la secuencia completa de la región codificante de los genes mediante la técnica RACE PCR.

3.2.10.2 RACE

Las secuencias completas de la región codificante de PAs de *G. verum* y *F. carica* fueron obtenidas mediante la técnica RACE PCR (“Rapid Amplification of cDNA Ends”) (Chenchik et al. 1996). Esta técnica permite amplificar los extremos 5’ y 3’ desconocidos de un ADNc utilizando la información de una secuencia conocida dentro del transcrito (esquema de la técnica en Figura 16). Se empleó el kit Marathon® cDNA Amplification Kit (Clontech, Mountain View, CA, USA), siguiendo el manual suministrado por la casa comercial (www.clontech.com, protocolo N° PT1115-1), que se resume a continuación.

3.2.10.2.1 Síntesis de ADNc

La síntesis de ADNc se realizó utilizando la enzima transcriptasa inversa AMV y el cebador cDNA Synthesis Primer (Tabla 10) siguiendo el protocolo indicado en el apartado 3.2.7.1.

3.2.10.2.2 Síntesis de la segunda hebra de ADNc

La segunda hebra de ADNc se sintetizó combinando 10 µL de la reacción obtenida en el apartado anterior con dNTPs 0.2 mM (cada

uno) y un cóctel enzimático constituido por 24 U ADN polimerasa I de *E. coli*, 4.8 U ADN ligasa de *E. coli* y 1 U RNasa H. La reacción se llevó a cabo en tampón KCl 100 mM, sulfato amónico 10 mM, MgCl₂ 5 mM, β-NAD 0.15 mM, Tris 20 mM (pH 7.5) y 0.05 mg/mL BSA en un volumen final de 80 μL y se incubó durante 1.5 horas a 16 °C. La reacción se detuvo añadiendo 4 μL de una mezcla EDTA 0.2 M / glucógeno (2mg/mL) y se purificó el ADN mediante el método de fenol:cloroformo y precipitación con etanol en presencia de acetato amónico. El ADN de doble hebra obtenido se resuspendió en 10 μL de H₂O_{mQ}.

3.2.10.2.3 Ligación del adaptador al ADN doble hebra

El ADN de doble hebra obtenido en el apartado anterior fue ligado al adaptador suministrado por el kit. La mezcla de reacción contenía: 5 μL de ADN doble hebra, Marathon cDNA Adaptador 2 μM, 400 U de T4 ADN ligasa, tampón de reacción Tris 50 mM pH 7.8, MgCl₂ 10 mM, DTT 1 mM, ATP 1 mM, polietilenglicol 5 % (p/v) (MW 8000) en un volumen final de 10 μL. La ligación se llevó a cabo a 16 °C durante toda la noche, tras lo cual se incubó la mezcla de reacción a 70 °C durante 5 min para inactivar la ligasa. El ADN fue cuantificado y almacenado a -20 °C. Cuando fue necesario el ADN se diluyó 1/10 en tampón Tricina-KOH 10 mM pH 8.5, EDTA 0.1 mM antes de su uso.

3.2.10.2.4 RACE PCR

El ADN unido al adaptador se utilizó como molde para realizar la 5' y 3' RACE PCR (Figura 16).

El adaptador suministrado por el kit no contiene el sitio de unión para el cebador AP1; durante la primera ronda de la RACE PCR el cebador específico es extendido hasta el final del adaptador, creando el sitio de unión para el cebador AP1 en los siguientes ciclos de la PCR. Este diseño del adaptador reduce los productos de PCR inespecíficos.

En cada una de las reacciones de PCR se utilizaron como cebadores el AP1 primer (cebador que hibrida en el adaptador) y un

cebador específico diseñado en base a las secuencias internas de ~1 kb obtenidas a partir de las PCRs realizadas con la pareja de cebadores degenerados pAPF y pAPR (apartado 3.2.10.1). La secuencia de los cebadores específicos diseñados en base a cuatro de las secuencias parciales obtenidas pueden consultarse en la Tabla 10, en la que figuran con los nombres pMgv1r (5'RACE) y pMgv1f (3'RACE) diseñados para obtener la secuencia codificante del gen que denominaremos *gv1AP*, pMgv2r (5'RACE) y pMgv2f (3'RACE) para obtener la secuencia del gen *gv2AP*, pMfc1r (5'RACE) y pMfc1f (3'RACE) para el gen *fc1AP*, y pMfc2r (5'RACE) y pMfc2f (3'RACE) para el gen *fc2AP*. Los cebadores específicos fueron diseñados de forma que los productos de la 5' y 3' RACE PCR se solapasen, facilitando la obtención de la secuencia completa del gen.

Las reacciones de amplificación se realizaron utilizando la polimerasa Accuzyme y las condiciones fueron:

(94 °C 30 s)_{1x} (94 °C 5 s, 72 °C 3 min)_{30x} para muestras procedentes de *G. verum*.

(94 °C 30 s)_{1x} (94 °C 5 s, 68 °C 2 min)_{30x} para muestras procedentes de *F. carica*.

Los productos de PCR (4 µL) fueron clonados en el vector pCR®-Blunt II-TOPO®, el ADN plasmídico purificado y enviado a secuencia utilizando los cebadores F17M13 y R19M13.

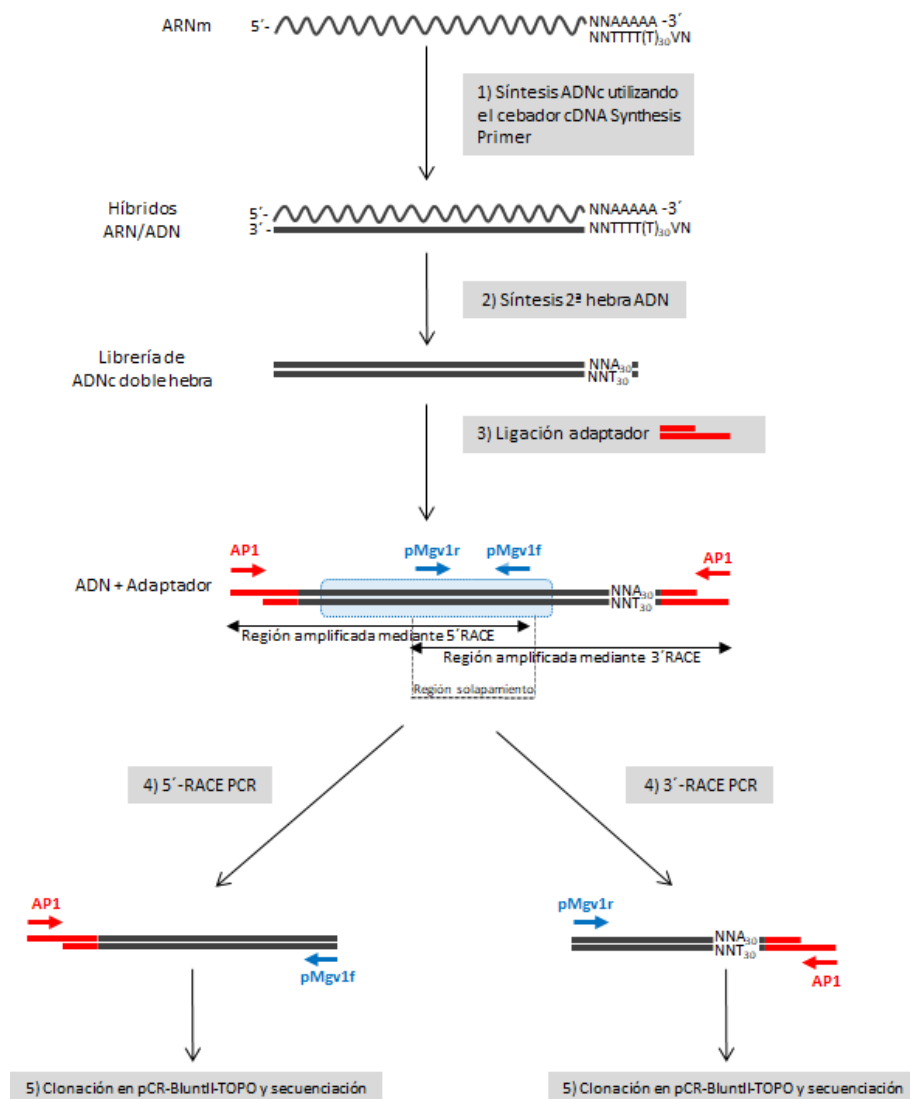


Figura 16. Esquema de la técnica RACE PCR para la obtención de la secuencia codificante del gen *gv1AP*. El recuadro azul indica la secuencia interna, determinada mediante los cebadores degenerados pAPF y pAPR, utilizada para diseñar los cebadores específicos pMgv1r y pMgv1f. Las secuencias obtenidas a partir de la 5' y 3' RACE PCR se solapan.

3.2.10.3 Amplificación del gen *gvIAP* desde ADN genómico

La amplificación parcial del ADN genómico correspondiente al gen *gvIAP* se realizó utilizando las parejas de cebadores pGv1-EcoRI/pMgv1r y pMgv1r/pGv1-XbaI (Tabla 10) mediante la polimerasa PrimeSTAR® HS DNA Polymerase (Takara).

El programa de amplificación empleado fue:

(98 °C 3 min)_{1x} (98 °C 10 s, 68 °C 5 min)_{30x} (68 °C 7 min)_{1x}

Los productos de estas amplificaciones fueron analizados mediante electroforesis en gel de agarosa al 1 % (p/v) y las bandas purificadas (apartado 3.2.1.6). El producto de PCR se fosforiló utilizando la enzima T4 polinucleótido quinasa (Takara).

El vector pUC19 fue digerido con la enzima SmaI y purificado (apartado 3.2.1.7). El vector lineal fue desfosforilado utilizando la enzima Shrimp Alkaline Phosphatase (SAP) (Promega), tras lo cual fue inactivada (65 °C 15 min).

Los extremos romos de vector e inserto (en relación 1:3) fueron ligados empleando la enzima ADN ligasa T4 (Takara). La mezcla de ligación fue empleada para transformar células quimiocompetentes de *E. coli* (apartado 3.3.1.4). Se extrajo el ADN plasmídico desde varios clones recombinantes y se secuenció utilizando los cebadores F17M13 y R19M13 (Tabla 10).

3.3 Generación de microorganismos genéticamente modificados

3.3.1 Obtención de cepas de *E. coli* recombinantes

3.3.1.1 Construcción del vector pET30a-progv1AP

El vector pET30a-progv1AP se construyó para la expresión de la preprogaline A en *E. coli* cepa BL21(DE3).

El **inserto** conteniendo la región que codifica para la preprogaline A sin el péptido señal, que denominaremos progaline A, se obtuvo mediante PCR realizada con la enzima Accuzyme utilizando como

molde ADNc de *G. verum* y los cebadores pGv1-EcoRI/pGv1-XbaI (Tabla 10). El programa empleado en el termociclador fue: (95 °C 30 s, 61 °C 30 s, 68 °C 2 min)_{30x} (68°C 7 min)_{1x}.

El producto de PCR se sometió a electroforesis en gel de agarosa y la banda obtenida fue purificada mediante el kit GEHealthcare y clonada en el vector pCR®-Blunt II-TOPO®. El ADN plasmídico (denominado TOPO-progv1AP) se extrajo con el kit Plasmid Midi Kit (Qiagen) y se cortó con la enzima de restricción EcoRI.

El **vector pET30** (Figura 12) se cortó con la misma enzima EcoRI, se defosforiló con SAP, se sometió a electroforesis en gel de agarosa 0.9 % (p/v) y se purificó (apartado 3.2.1.6).

El inserto y el vector fueron ligados utilizando la enzima T4 DNA ligasa (Takara) y las relaciones molares 1:1 y 1:3.

3.3.1.2 Construcción de los vectores pET21a-progv1AP y pET21a-profc1AP

Los vectores pET21a-progv1AP y pET21a-profc1AP se construyeron para la expresión de la preprogaline A y preprofigain A respectivamente en *E. coli* cepa Rosetta-gami (DE3).

Los **insertos** conteniendo la región que codifica para la preprogaline A o la preprofigain A, sin el péptido señal, que denominaremos progaline A y profigain A respectivamente, se obtuvieron mediante PCR realizada con la enzima Accuzyme utilizando como molde ADNc de *G. verum* y los cebadores F-NheIprogv1/R-NheIprogv1 o el ADNc de *F. carica* y los cebadores F-NheIprofc1/R-NheIprofc1. Estos cebadores (Tabla 10) introducen el sitio de restricción NheI en posición 5' y 3'. El programa empleado en el termociclador fue: 95 °C, 30 s; 55 °C 30 s; 68 °C 2 min)_{30x} (68 °C 7 min)_{1x}.

El producto de PCR se sometió a electroforesis en gel de agarosa y la banda obtenida fue purificada (apartado 3.2.1.6) y cortada con la enzima NheI.

El **vector pET21a(+)** (Figura 13) se cortó con la misma enzima, NheI, se defosforiló con SAP, se sometió a electroforesis en gel de agarosa y se purificó.

El inserto y el vector fueron ligados utilizando la enzima T4 ADN ligasa (Takara).

3.3.1.3 Análisis de las construcciones

Las construcciones obtenidas en los apartados 3.3.1.1 y 3.3.1.2, que denominamos pET30a-progv1AP, pET21a-progv1AP y pET21a-profc1AP se clonaron en primer lugar en la cepa de *E. coli* DH5 α , para poder confirmar las secuencia eran las adecuadas. La clonación en esta cepa, que no contiene el gen de la T7 ARN polimerasa, antes de abordar la transformación de la cepa BL21(DE3) o Rosetta-gami (DE3) se realizó para evitar la inestabilidad plasmídica derivada de la producción de proteínas potencialmente tóxicas en estos hospedadores.

El ADN plasmídico fue extraído y analizado con las enzimas de restricción: HindIII en el caso de la construcción pET30a-progv1AP; PstI en el caso de la construcción pET21a-profc1AP y BamHI cuando la construcción fue pET21a-progv1AP.

Tras verificar que la secuencia nucleotídica era correcta utilizando los cebadores T7promotor y T7terminador (Tabla 10), se realizó la transformación de la cepa de expresión BL21(DE3) con la construcción pET30a-progv1AP o de la cepa Rosetta-gami (DE3) con las construcciones pET21a-progv1AP o pET21a-profc1AP.

3.3.1.4 Transformación de *E. coli*

3.3.1.4.1 Preparación de células competentes

Se inocularon 5 mL de medio LB con una colonia de la cepa de *E. coli* a transformar y se cultivó a 37 °C y 180 rpm durante toda una noche, hasta alcanzar la fase estacionaria. Se utilizaron 2 mL de este cultivo para inocular 100 mL de medio LB y se incubó en agitador orbital a 37 °C y 180 rpm hasta alcanzar una DO₅₅₀ de 0.3-0.4. El cultivo se enfrió sobre hielo y se centrifugó a 3000 g y 4 °C durante 10 min. Las células se resuspendieron en 30 mL de una solución de TfbI estéril y fría. Se realizó una centrifugación en las mismas condiciones descritas anteriormente y las células fueron resuspendidas en 4 mL de TfbII estéril. La suspensión así obtenida se repartió en viales en alícuotas de 100 µL y fueron congeladas inmediatamente en nitrógeno líquido. Las células quimiocompetentes se almacenaron a -70 °C hasta su uso.

La composición y preparación de las soluciones empleadas fue:

- **Solución TfbI (500 mL):** 1.47 g de CH₃COOK, 5.95 g de MnCl₂, 6.05 de RbCl₂, 0.74 g de CaCl₂-6H₂O, 75 mL de glicerol. Se ajustó a pH 5.8 con ácido acético glacial 0.2 M, se esterilizó mediante un filtro de 0.22 µm y almacenó a 4 °C.

- **Solución TfbII (100mL):** 10 mL de un *stock de MOPS, 1.6 g de CaCl₂-6H₂O, 0.12 g de RbCl₂, 18 mL de glicerol. La solución se esterilizó por filtración (0.22 µm) y almacenó a 4 °C en oscuridad.

*Stock de MOPS (100mL): 2.09 g de MOPS, 0.34 g de CH₃COONa, 0.18 g de EDTA. La solución se ajustó a pH 7 y se esterilizó por filtración (0.22 µm).

3.3.1.4.2 Transformación mediante choque térmico

Las células quimiocompetentes (100 µL) se descongelaron en hielo y se mezclaron con el ADN plasmídico o con la mezcla de ligación (50 ng en un volumen máximo de 10 µL). El vial se mantuvo

en hielo durante 30 min, transcurridos los cuales se aplicó un choque térmico sumergiendo las células en un baño a 42 °C durante 45 s y transfiriéndolas a hielo durante 2 min. Se añadieron 800 µL de medio LB o medio SOC y la mezcla se transfirió a un tubo de 10 mL, que se incubó durante 45 min a 37 °C y 180 rpm para permitir la expresión génica de la resistencia empleada para la selección. Finalmente, las células se sembraron en medio sólido LB suplementado con el antibiótico cuya resistencia porta el plásmido empleado en la transformación. El medio sólido se suplementó con X-Gal e IPTG cuando fue necesario. Las placas se incubaron a 37 °C.

Cuando se utilizaron las células quimiocompetentes One Shot®TOP10 *E. coli* (Invitrogen) se siguió el protocolo suministrado por el fabricante.

3.3.2 Obtención de cepas de *P. pastoris* recombinantes

3.3.2.1 Construcción de los vectores pPICZα-progv1AP, pPICZα-profc1AP y pPICZα-fc1AP

Los vectores pPICZα-progv1AP, pPICZα-profc1AP y pPICZα-fc1AP se construyeron para la expresión de la preprogaline A y la preprofigain A en *P. pastoris* bajo un promotor inducible por metanol.

El **inserto** que codifica para la preprogaline A sin el péptido señal, que denominamos progaine A, se obtuvo mediante corte del vector TOPO-progv1AP (apartado 3.3.1.1) con las enzimas de restricción EcoRI y XbaI.

Los **insertos** que codifican para la preprofigain A sin el péptido señal, que denominamos profigain A, y para la preprofigain A sin el péptido señal y sin el prosegmento, que denominaremos figain A, se obtuvieron mediante PCR realizada con la enzima PfuUltra® utilizando como molde ADNc obtenido desde el peciolo de *F. carica* y los cebadores Fc1-EcoRI/pFc1-XbaI para amplificar la figain A y

pFc1-EcoRI/pFc1-XbaI para la amplificación de la profigain A (Tabla 10). Estos cebadores introducen los sitios de restricción EcoRI y XbaI en posición 5' y 3' respectivamente del producto amplificado. El programa empleado en el termociclador fue: (95 °C 30 s, 55 °C 30 s, 72 °C 2 min)_{5x} (95 °C 30 s, 60 °C 30 s, 72 °C 2 min)_{30x} (72 °C 10 min)_{1x}.

Los productos de PCR se purificaron desde gel (apartado 3.2.1.6) y se clonaron en el vector pCR®-Blunt II TOPO®. Se realizó una midiprep a varios de los clones obtenidos y se analizó el ADN plasmídico mediante corte con las enzimas EcoRI y XbaI.

El **vector** pPICZα (Figura 14) se cortó con las enzimas de restricción EcoRI y XbaI, se purificó desde gel y se defosforiló con la enzima SAP.

Inserto y vector fueron ligados utilizando la enzima T4 ADN ligasa (Takara).

Las construcciones obtenidas se clonaron en una cepa de *E. coli* TOP10 y se seleccionaron en placas de LB suplementadas con 25 µg/mL de zeocina. Las construcciones extraídas desde *E. coli* fueron analizadas mediante corte con enzimas de restricción y secuenciadas utilizando los cebadores 5'AOX1 y 3'AOX1 (Tabla 10). Las construcciones con la secuencia correcta fueron utilizadas para la transformación de *P. pastoris* siguiendo la metodología indicada en el apartado 3.3.2.3.

3.3.2.2 Construcción de los vectores pGAPZα-progv1AP, pGAPZα-progv2AP, pGAPZα-profc1AP y pPGAPZα-fc1AP

Los vectores pGAPZα-progv1AP, pGAPZα-progv2AP, pGAPZα-profc1AP y pPGAPZα-fc1AP se construyeron para la expresión constitutiva de la preprogaline A, preprogaline B y preprofigain A en *P. pastoris*.

Los **insertos** que codifican para la progaine A, la profigain A y la figain A utilizados para la construcción de los vectores pGAPZ α -progv1AP, pGAPZ α -profc1AP y pPGAPZ α -fc1AP fueron los mismos que en el apartado 3.3.2.1.

El **inserto** que codifica para la preprogaine B sin la región que codifica para el péptido señal, utilizado para la construcción pGAPZ α -progv2AP se obtuvo mediante PCR utilizando como molde ADNc de *G. verum* y los cebadores pGv2-EcoRI y pGv2-XbaI (Tabla 10). Estos cebadores introducen los sitios de restricción EcoRI y XbaI en posición 5' y 3' respectivamente del producto amplificado. El programa empleado en el termociclador fue: (98 °C 10 s, 58 °C 5 s, 72 °C 1min)_{30x}

Los productos de PCR se purificaron desde gel (apartado 3.2.1.6) y se digirieron con las enzimas de restricción EcoRI y XbaI.

El **vector** pGAPZ α (Figura 15) se cortó con las mismas enzimas de restricción que el inserto correspondiente, se purificó desde gel y se defosforiló con la enzima SAP.

Los insertos y el vector fueron ligados utilizando la enzima T4 ADN ligasa (Takara).

Las construcciones obtenidas se clonaron en la cepa de *E. coli* TOP10 y se seleccionaron en placas de LB a pH 7.5 suplementadas con 25 μ g/mL de zeocina. Las construcciones fueron extraídas desde *E. coli*, analizadas mediante corte con enzimas de restricción y secuenciadas. Las construcciones con la secuencia correcta fueron utilizadas para la transformación de *P. pastoris* siguiendo la metodología indicada en el apartado 3.3.2.3.

3.3.2.3 Transformación de *P. pastoris*

3.3.2.3.1 Preparación de células competentes

P. pastoris se cultivó en 5 mL de YPD toda la noche a 30 °C con agitación. Este cultivo (0.1-0.5 mL) se utilizó para inocular 500 mL de YPD. Cuando se alcanzó una DO₆₀₀ entre 1.3-1.5 el cultivo se centrifugó a 1500 g durante 5 min a 4 °C y se lavó el pellet con agua destilada estéril fría (0 °C) dos veces, la primera con un volumen de 500 mL y la segunda con 250 mL. El pellet se resuspendió en 20 mL de sorbitol 1 M (0 °C) y se realizó una nueva centrifugación, para finalmente resuspender las células en 1 mL sorbitol 1 M (0 °C). Las células se mantuvieron en hielo hasta su uso.

3.3.2.3.2 Transformación mediante electroporación

Las construcciones plasmídicas fueron linearizadas empleando la enzima de restricción SacI cuando se utilizó el vector pPICZα, o con el enzima BspHI cuando el vector empleado fue pGAPZα. La linearización se realizó para facilitar los fenómenos de recombinación entre regiones homólogas existentes entre la construcción y el ADN genómico de *P. pastoris* X-33. Cuando las construcciones estaban completamente digeridas (comprobado mediante electroforesis) se inactivó el enzima de restricción y se realizó una purificación del ADN mediante extracción con fenol:cloroformo y precipitación con etanol, utilizando 1/10 volúmenes de acetato sódico 3 M y 2.5 volúmenes de etanol 100 %. La cantidad de ADN para realizar la transformación fue de 5-10 µg de ADN en un volumen final de 10 µL.

Las células de *P. pastoris* fueron transformadas mediante electroporación utilizando un aparato GenePulser II (Bio-Rad, Laboratories, Richmond, CA). Las células (80 µL) se mezclaron con 5-10 µg de ADN linearizado. La mezcla se transfirió a una cubeta de 0.2 cm (Bio-Rad) e incubó en hielo durante 5 min, transcurridos los cuales se aplicó un pulso eléctrico (1.8 kV, 25 µF, 200 Ω). Se añadió 1

mL de una solución 1 M sorbitol frío a la cubeta y se transfirió la mezcla a un tubo estéril de 15 mL, que se incubó a 30 °C sin agitación durante 3-5 horas. Las células transformadas se seleccionaron en placas de YPDS suplementadas con 100 µg/mL zeocina, incubadas a 30 °C durante 3-6 días. Se realizó una segunda ronda de selección con las colonias obtenidas.

Las colonias fueron analizadas para verificar que las construcciones habían sido integradas en el genoma. Se realizó una PCR utilizando como molde el ADN genómico extraído mediante el kit “MasterPure™Yeast DNA Purification kit” o bien directamente desde las colonias utilizando el reactivo “Lyse and Go PCR Reagent” (Thermo Scientific).

La búsqueda de clones con múltiples copias del casete de inserción se realizó en placas de YPDS suplementadas con una concentración de 2000 µg/mL de zeocina (Cregg et al. 2009).

3.4 Técnicas de análisis y caracterización de proteínas

3.4.1 Electroforesis de proteínas en geles de poliacrilamida

La separación de proteínas se realizó según el procedimiento de Laemmli (Laemmli 1970) empleando los soportes del equipo Mini Protean II (BioRad). Los geles se realizaron en un sistema discontinuo, compuestos de un gel concentrador superior al 4 % y un gel separador inferior al 10 % o 12 %, utilizando una solución Acrylamide/Bis-acrylamide 30 %, 29:1 (Bio-Rad, China). Las muestras se disolvieron (1:4) en tampón de carga 4X (azul de bromofenol 0.04 %, Tris-HCl 240 mM, pH 6.8, glicerol 40 %, SDS 8 %) al que se añadió β-ME 10 % (14.4 M) cuando se quisieron romper los puentes disulfuro de las muestras (por defecto las proteínas fueron analizadas en condiciones reductoras a no ser que se especifique lo contrario). Las muestras se calentaron a 95°C durante 7 minutos. La electroforesis se desarrolló en tampón Tris glicina (Tris 25 mM,

glicina 192 mM y SDS 0.1 %) aplicando una corriente de 60 V, hasta que el frente de migración alcanzó la frontera del gel separador, tras lo cual se subió el voltaje a 120 V. En el caso del ensayo para determinar la hidrólisis de las caseínas el tampón utilizado fue tampón Tris-tricina (Tris 100 mM, tricina 100 mM, SDS 0.1 %).

Las proteínas se detectaron mediante:

- i) tinción azul de Coomassie: las proteínas se fijaron con una solución 50 % metanol, 7 % ácido acético durante 15 min y se tiñeron con azul de Coomassie [azul de Coomassie R-250 0.1 % (Bio-Rad), metanol 40 % (v/v), ácido acético 10 % (v/v)] durante 1-2 horas. Para eliminar el fondo el gel se sumergió en una solución decolorante [metanol 40 % (v/v), ácido acético 10 % (v/v)], tras lo cual se transfirió a una solución ácido acético 10 %.
- ii) tinción con sales de plata según el procedimiento descrito por Gromova and Celis (2006).

El peso molecular de las proteínas se estimó por comparación con el patrón de migración de las bandas del marcador de pesos moleculares Precision Plus Protein Standard (BioRad), utilizando el Sistema de documentación de geles ChemiDoc XRS+ (BioRad) y el software Bio-Rad Image Lab™.

3.4.2 Cuantificación de proteínas

La concentración de proteínas se determinó utilizando el Pierce™ Coomassie (Bradford) Protein Assay Kit (Thermo Scientific, Rockland, IL, USA) basado en el método de Bradford (1976); la recta patrón se realizó utilizando BSA.

El volumen de muestra para el análisis de las diferentes fracciones celulares obtenidas desde las células de *E. coli* transformadas con el sistema pET se normalizó para que la intensidad de las bandas teñidas con Coomassie reflejase la cantidad relativa de cada una de las fracciones, como se indica en el manual pET System Manual (11th Edition, Novagen).

En otras ocasiones la cuantificación se realizó desde gel, comparando la intensidad de la banda de la proteína con la de bandas de concentración conocida del marcador de pesos moleculares, utilizando el sistema de documentación de geles ChemiDoc XRS+(BioRad) y el software Bio-Rad Image Lab™.

3.4.3 Western blot

Un péptido sintético, con la secuencia correspondiente a los aminoácidos 243 a 255 de la preprogalina B (PSHFKGEHVYAKV) fue químicamente sintetizado y utilizado para la producción de anticuerpos policlonales en conejo, utilizando KLH (keyhole limpet hemocyanin) como “carrier”. El suero conteniendo los anticuerpos, que denominaremos PV13, fue producido por la empresa Biomedal S.L. (Sevilla, España) y utilizado para realizar el Western blot. La técnica Western blot se realizó según describen Mahmood y Yang (2012). Tras la electroforesis, las proteínas fueron transferidas a 100 V durante 1 h 15 min a una membrana de fluoruro de polivinilideno (PVDF; membrana de transferencia Immobilon-P, Millipore, Billerica, USA) con diámetro de poro 0.45 μm , en una célula de transferencia Mini Trans-Blot® Cell (Bio-Rad), utilizando un tampón que contenía Tris 25 mM, glicina 192 mM, metanol 10 % (v/v). Una vez realizada la transferencia la membrana se lavó con tampón TBS (Tris 50 mM, NaCl 150 mM, pH 7.5) y se incubó durante 1 hora en solución bloqueante al 1 % (Western Blocking Reagent; Roche, Indianapolis, USA). A continuación la membrana se incubó con el suero PV13 diluido 1:1000 en solución bloqueante 0.5 % durante 1 hora, transcurrida la cual, se realizaron varios lavados con TBS suplementado con Tween 20 0.1 % (TBST) (Millipore, Germany). Tras los lavados se realizó una incubación con el anticuerpo secundario anti-conejo conjugado a peroxidasa de rábano (HRP, horseradish peroxidase) (Goat Anti-Rabbit IgG (H+L)-HRP

Conjugate, Bio-Rad) diluido 1:10000 en solución bloqueante 0.5 %. Tras varios lavados con TBST las proteínas se visualizaron mediante quimioluminiscencia empleando el sustrato Clarity™ Western ECL Blotting Substrate (Bio-Rad) en un equipo ChemiDoc XRS+ (Bio-Rad) y el software Bio-Rad Image Lab™.

El peso molecular de las proteínas se estimó utilizando un marcador de proteínas biotiniladas (Biotinylated Molecular Weight Marker, Sigma-Aldrich, USA), que fueron detectadas añadiendo un conjugado extravidina peroxidasa (Sigma-Aldrich, USA) en una dilución 1:5000. La adición del conjugado se realizó al mismo tiempo que el anticuerpo secundario.

3.4.4 Extracción de proteínas

3.4.4.1 Extracción de proteínas desde cultivos de *E. coli*

La obtención de las proteínas desde diferentes fracciones celulares de las cepas de *E. coli* BL21(DE3) o Rosetta-Gami (DE3) pET se realizó siguiendo los protocolos detallados en el pET System Manual (11th Edition, Novagen).

➤ Extracción del contenido proteico total

Las células desde 1mL de cultivo obtenidas mediante centrifugación (10000 g 1 min) fueron resuspendidas en 100 µL de tampón fosfato salino (PBS: KH₂PO₄ 1.5 mM, Na₂HPO₄ 8.1 mM, NaCl 0.14 M, KCl 2.7 mM, pH 7.4). Se añadieron 100 µL de tampón de carga de proteínas (4X) y la muestra se sometió a ultrasonidos a un nivel entre 2–3, ciclos de trabajo 20–30 % durante 8–10 ráfagas en un sonicador (Branson Sonifier 250, Connecticut, USA). La muestra se calentó 7 min a 95 °C y almacenó a -20 °C hasta su análisis mediante SDS-PAGE.

➤ Extracción de la fracción citoplasmática soluble

Las células se obtuvieron mediante centrifugación (10000 g, 1 min) y se determinó el peso húmedo de las mismas. Se sometieron a un ciclo de congelación/descongelación y se resuspendieron en tampón de lisis (Tris-HCl 50 mM, glicerol 5 %, NaCl 50 mM, EDTA 2 mM, PMSF 1 mM, pH 8) utilizando 7 mL de tampón por gramo de células; se añadieron 100 µg de lisozima, MgCl₂ (c.s.p. 2 mM) y 100 µg DNasa por cada mL de tampón de lisis. La solución se mantuvo en agitación hasta la rotura celular y mediante centrifugación a 4 °C (16000 g, 20 min) se separó la fracción citoplasmática insoluble (pellet) y la fracción citoplasmática soluble (sobrenadante). El sobrenadante se combinó con tampón de carga de proteínas, se calentó a 95 °C 7 min y almacenó a -20 °C hasta su análisis.

Se realizó una modificación del proceso suplementando el tampón de lisis con el detergente CHAPS a una concentración 10 mM, e introduciendo un paso de sonicación tras la rotura celular.

➤ **Extracción de la fracción citoplasmática insoluble**

- Muestras destinadas a ser analizadas mediante SDS-PAGE:

La fracción citoplasmática insoluble obtenida durante la extracción de la fracción citoplasmática soluble (apartado anterior) se lavó con 750 µL de una solución Tris HCl 20 mM pH 7.5 por cada 40 mL de cultivo inicial y se resuspendió en 1.5 mL de SDS 1 % mediante sonicación. Se combinaron 100 µL de muestra con 100 µL de tampón de carga de proteínas 4X, se calentó a 95 °C y almacenó a -20 °C hasta su análisis.

- Muestras destinadas a ser purificadas mediante cromatografía de afinidad a iones inmovilizados (IMAC):

La fracción citoplasmática insoluble obtenida desde la fracción citoplasmática soluble se sometió a ultrasonidos utilizando una lanza de sonicación (Branson sonifier 250, USA), el proceso se realizó sobre

hielo. La suspensión se centrifugó a 12000 g durante 20 min a 4 °C y el pellet se resuspendió en PBS / urea 8M, pH 7.5 en un ratio 1:10 (p/v) y sometió a un nuevo ciclo de ultrasonidos.

3.4.4.2 Extracción de proteínas desde cultivos de *P. pastoris*

➤ Extracción desde el interior celular

Las células obtenidas mediante centrifugación a 4 °C se lavaron y resuspendieron en tampón de lisis frío (fosfato sódico 50 mM pH 7.4, PMSF 1 mM, EDTA 1 mM, glicerol 5 %) a una DO₆₀₀ de 50-100. Se añadió un volumen de perlas de vidrio de 0.5 mm y las células se rompieron mediante agitación vigorosa durante 30 segundos e incubación en hielo otros 30 segundos, durante 8 ciclos. La muestra se centrifugó a 4 °C y 12000 g durante 10 min. El sobrenadante se transfirió a un nuevo vial y se analizó mediante SDS-PAGE.

➤ Sobrenadantes

Los cultivos celulares se centrifugaron a 5000 g durante 15 min a 4 °C y el sobrenadante se filtró a través de un filtro de 0.22 µm.

Las muestras fueron analizadas en el momento o congeladas en N₂ líquido y almacenadas a -80 °C hasta su uso.

Cuando fue necesario se realizó la precipitación mediante TCA de las proteínas obtenidas para concentrarlas (Koontz 2014) o bien se dializaron y concentraron mediante membranas o tubos de ultrafiltración de NMWCO de 10 o 30 kDa (Millipore, Bedford, USA).

3.4.5 Purificación de las proteasas

3.4.5.1 Purificación de las proteasas desde *E. coli*

La purificación de los cuerpos de inclusión desde *E. coli* se realizó a pequeña escala, en viales de 1.5 mL, mediante cromatografía

de afinidad, empleando los tampones indicados en el manual Profinity IMAC Resins (BioRad) en condiciones desnaturalizantes.

La resina (500 μ L) (Profinity TMIMAC Ni-Charged Resin, Bio-Rad, CA) se centrifugó a 5000 g durante 1 min y se realizaron dos lavados con tampón de unión (fosfato sódico 50 mM, NaCl 300 mM, imidazol 5 mM, urea 7 M, pH 8). Se mezclaron 900 μ L de la resina disuelta en tampón de unión con 100 μ L de la fracción citoplasmática insoluble disuelta en PBS /urea 8 M pH 7.5 (apartado 3.4.4.1). La mezcla se mantuvo en agitador orbital a temperatura ambiente durante 10 min, transcurridos los cuales se realizó una centrifugación a 3000 g durante 1 min. El pellet se resuspendió en 500 μ L de tampón de lavado (fosfato sódico 50 mM, NaCl 300 mM, imidazol 10 mM, urea 7 M, pH 8) y se realizó una nueva centrifugación a 3000 g durante 1 min. Se añadieron 200 μ L de tampón de elución a la resina unida a la proteína (fosfato sódico 50 mM, NaCl 300 mM, imidazol 0.25 mM, urea 7 M, pH 8) y se mantuvo en agitador orbital durante 5 min a temperatura ambiente. Se realizó una centrifugación a 3000 g durante 1 min y se transfirió el sobrenadante conteniendo la proteína recombinante a un nuevo vial. Se realizó una segunda elución en las mismas condiciones que la primera.

Se ensayaron tres concentraciones de imidazol en el tampón de elución: 0.25 M, 0.50 M y 0.75 M. La concentración de imidazol utilizada en el resto de los análisis fue de 0.25 M.

3.4.5.2 Purificación de la progaline B desde *P. pastoris*.

La purificación de la proteasa recombinante se llevó a cabo mediante cromatografía de intercambio iónico.

Transcurridos tres días de cultivo el medio extracelular (500 mL) de un cultivo de *P. pastoris* transformado con la construcción pGAPZ α -progv2AP fue separado mediante centrifugación, filtrado a través de 0.22 μ m, concentrado y dializado frente a Tris-HCl 20 mM pH 7.5 (tampón A) mediante una unidad de ultrafiltración (Stirred

Ultrafiltration Cell, Millipore Amicon) equipado con una membrana de celulosa regenerada YM10, de peso molecular de corte 10 kDa (Millipore, Bedford, USA). El sobrenadante dializado se aplicó a una columna (1.5 x 10 cm) DEAE Bio-Gel A (Bio-Rad, Hercules, USA) previamente equilibrada con tampón A. La columna se lavó con tampón A suplementado con NaCl 50 mM y la proteasa se eluyó con el mismo tampón suplementado con NaCl 300 mM, recogién dose fracciones de 5 mL de volumen. Este proceso se llevó a cabo en una cámara a una temperatura de 4 °C. Las fracciones fueron analizadas para determinar su actividad proteolítica utilizando FTC- κ -caseína. Las fracciones que presentaban una mayor actividad fueron analizadas mediante SDS-PAGE al 12 % y aquellas en las cuales la progaline B era la proteína mayoritaria fueron seleccionadas para el siguiente paso de purificación. Las fracciones seleccionadas se agruparon, se dializaron frente a tampón A y aplicaron a una columna FPLC monoQ (HR 5/5 Amersham Biosciences). La elución se realizó con cantidades crecientes de NaCl en el rango de 100 a 450 mM. Aquellas fracciones que presentaban actividad fueron dializadas frente a tampón fosfato sódico 20 mM pH 7.5 y analizadas mediante SDS PAGE 15% y tinción con sales de plata.

3.4.6 Renaturalización de los cuerpos de inclusión desde *E. coli*

3.4.6.1 Método de dilución rápida

A los cuerpos de inclusión purificados con la resina de níquel (IMAC) se les añadió β -ME a una concentración 100 mM. Las proteínas fueron replegadas mediante dilución (40 veces) en Tris base 20 mM (Castanheira et. al 2005). Las proteínas se añadieron lentamente a la solución en agitación, mediante un gotero, y posteriormente el pH se ajustó a pH 8, bajando 0.2 puntos cada 12 horas. Se probaron dos temperaturas para realizar el proceso de renaturalización: 4 °C y temperatura ambiente. Las muestras fueron concentradas y dializadas frente a Tris HCl 20 mM pH 8.

Las muestras replegadas fueron incubadas a 37 °C con un volumen de tampón a diferentes valores de pH para activarlas.

3.4.6.2 Método de plegado universal

Los cuerpos de inclusión se añadieron a una solución conteniendo los reactivos necesarios para que la concentración final fuese urea 8 M, Tris 0.1 M, glicina 1 mM, EDTA 1mM, β -ME 10mM, DTT 10 mM, glutatión 1 M en su forma reducida (GSH), glutatión 0.1 mM en su forma oxidada (GSSH), pH 10. La absorbancia a 280 nm de la solución se ajustó a 5 y se mantuvo 1 hora a temperatura ambiente. La muestra se diluyó lentamente en 20 volúmenes de Tris base 20 mM. Se ajustó el pH de la solución resultante a pH 9 con HCl y se mantuvo a 4 °C durante 24 horas. El pH se ajustó lentamente a pH 8, bajando 0.2 puntos cada 12 horas, el proceso se realizó en una cámara a 4 °C. Las muestras fueron concentradas y dializadas frente a Tris HCl 20 mM pH 8 conteniendo urea 0.4 M .

Las muestras replegadas fueron incubadas a 37 °C durante toda la noche con tampones a diferentes valores de pH para activarlas.

3.4.7 Espectrometría de masas

Los polipéptidos destinados a ser analizados mediante espectrometría de masas fueron escindidos desde un gel SDS-PAGE 12 % teñido con azul de Coomassie mediante un escalpelo, en tiras de aproximadamente 1 x 3 mm, y digeridos con tripsina utilizando el kit In-Gel DigestZP (Millipore). Los péptidos fueron analizados en un espectrómetro de masas Bruker Reflex III MALDI-TOF y en un nanocromatógrafo líquido (Easy-nLC Proxeon) acoplado a un espectrómetro de masas de trampa de iones, modelo Amazon ETD (Bruker Daltonics, Bremen, Germany). Los análisis se llevaron a cabo en el servicio de Espectrometría de Masas y Proteómica de la RIAIDT (Rede de Infraestruturas de Apoio á Investigación e ó Desenvolvemento Tecnolóxico) de la Universidad de Santiago de Compostela.

3.4.8 Deglicosilación enzimática

El sobrenadante de un cultivo crecido en YPD durante 72 horas de la cepa *P. pastoris* expresando la progaline B fue concentrado mediante ultrafiltración utilizando membranas YM10 (Millipore). La digestión de los oligosacáridos unidos mediante enlaces N-glicosídicos se llevó a cabo en un volumen final de 50 μ L conteniendo 10 μ L del sobrenadante concentrado (20 X) y 2 μ L de la enzima endoglicosidasa H de *Streptomyces griseus* (Sigma-Aldrich) en tampón fosfato sódico 50 mM pH 5. Como control negativo se utilizó agua destilada en lugar de enzima. La reacción se llevó a cabo a 37 °C durante 3 horas, y se detuvo mediante adición de tampón de carga del SDS-PAGE, con o sin presencia del agente reductor β -ME, las muestras se calentaron a 95 °C durante 7 min y se analizaron mediante Western blot.

3.5 Análisis informáticos

3.5.1 Análisis de secuencias de nucleótidos y aminoácidos

La base de datos utilizada para la búsqueda de secuencias genéticas fue Genbank, a través de la página web del NCBI (National Center for Biotechnology Information) (<http://www.ncbi.nih.gov/>).

Otras bases de datos de proteínas consultadas fueron:

- RCSB PDB (RCSB Protein Data Bank): base de datos que recoge la estructura tridimensional de proteínas y ácidos nucleicos (<https://www.rcsb.org/pdb/home/home.do>).
- MEROPS: base de datos de proteasas y sus inhibidores (Rawlings et al. 2016) (<https://www.ebi.ac.uk/merops/>).
- Enzyme Database BRENDA (BRaunschweig ENzyme DAtabase): base de datos de información molecular y bioquímica sobre enzimas y rutas metabólicas (<https://www.brenda-enzymes.org/>).

- UniProt: base de datos de proteínas, unión de las tres bases de datos Swiss-prot, TrEMBL y PIR-PSD (<http://www.uniprot.org/>).
- CDD (Conserved Domain Database): base de datos de dominios conservados del NCBI (Marchler-Bauer et al. 2017).

El software Vector NTI Advance (Invitrogen) fue utilizado para el análisis y alineamiento de las secuencias de nucleótidos y aminoácidos, así como para el diseño de cebadores.

La predicción de parámetros físicos se realizó mediante el programa ProtParam del servidor ExPASy (<http://www.expasy.org/tools/protparam.html>).

La detección de péptidos señal y sitios de glicosilación se realizó con los programas de predicción SignalIP 4.1 (<http://www.cbs.dtu.dk/services/SignalIP-4.1/>) y NetNGlyc 1.0 (<http://www.cbs.dtu.dk/services/NetNGlyc/>) del servidor CBS (Center for Biological Sequence Analysis).

La búsqueda de secuencias homólogas se realizó a través de la herramienta BLAST (Basic Local Alignment Search Tool) del NCBI.

La búsqueda de dominios conservados se realizó a través de la herramienta CD-Search del NCBI.

El árbol de las relaciones filogenéticas fue realizado utilizando el método de máxima verosimilitud (Hall 2013) y el software bioinformático MEGA 6 (Molecular Evolutionary Genetics Analysis) (Tamura et al. 2013). Las secuencias de aminoácidos de la preprogaline A, preprogaline B y preprofigain A fueron comparadas con 18 secuencias homólogas de PAs de plantas extraídas de la base de datos Genbank. El algoritmo seleccionado para realizar el alineamiento fue MUSCLE (Edgar 2004). Para la reconstrucción filogenética mediante MV, primero se realizó una búsqueda para

determinar los modelos evolutivos que se ajustaban mejor a los datos según su valor BIC (Bayesian Information Criterion). La estimación de la fiabilidad del árbol construido se hizo mediante un análisis *bootstrap* con 500 repeticiones (Felsenstein 1985).

El estudio de modelización de la estructura tridimensional se realizó utilizando el programa CPH-models-3.0 (Nielsen et al. 2010) y la visualización con el programa Pymol (Delano Scientific LLC 2006). Las estructuras 3D obtenidas fueron validadas mediante el programa ProSAweb (<https://prosa.services.came.sbg.ac.at>). (Wiederstein y Sippl 2007).

La predicción de la frecuencia de codones del gen que codifica para la proteína recombinante progaline B fusionada al factor- α en *P. pastoris* se llevó a cabo mediante Graphical Codon Usage Analyser (<http://gcu.schoedl.de/>).

El programa NetGene2 Server (<http://www.cbs.dtu.dk/services/NetGene2/>) se utilizó para facilitar la predicción de la localización de los intrones del gen *gvIAP* realizada con el software Vector NTI Advance (Invitrogen)

3.5.2 Análisis estadísticos

El tratamiento estadístico de los datos se realizó con el programa informático IBM SPSS Statistics 20; para la comparación entre dos grupos se empleó el test de t de Student y para la comparación entre más de dos grupos se empleó el análisis de la varianza de una vía (ANOVA).

3.6 Ensayos de actividad enzimática

3.6.1 Efecto de inhibidores sobre la actividad enzimática

Los extractos vegetales y las muestras enzimáticas se sometieron a estudios de inhibición utilizando los siguientes inhibidores de proteasas a la concentración final indicada: PMSF (1mM) (Sigma), E-64 (10 μ M) (Sigma) y pepstatina A (1.5 μ M) (Sigma). La pepstatina A se disolvió en metanol: acético (9:1), E-64 en agua y PMSF en etanol. Las muestras se incubaron durante 30 min a 37 °C en presencia de los inhibidores. En todas las pruebas se realizó un control de vehículo para comprobar que el disolvente no afectaba a la actividad enzimática.

3.6.2 Activación de las proteasas recombinantes

La activación de las proteasas recombinantes se realizó combinando las muestras enzimáticas con un volumen de tampón 0.2 M. Los tampones utilizados fueron tampón citrato en el rango de pH 2-3.5, tampón acetato sódico en el rango de pH 3.6-5.5, tampón fosfato sódico en el rango de pH 6-7 y Tris HCl en el rango 7-8.

3.6.3. Zimograma

El sobrenadante (25 μ l) concentrado (15 X) de un cultivo de *P. pastoris* expresando la progaline B se analizó en un gel de poliacrilamida 12 % copolimerizado con caseína (Ready Gel® Zymogram Precast Gels, Bio-Rad) siguiendo las instrucciones del fabricante. Las proteasas contenidas en el gel fueron renaturalizadas con una solución 2.5 % de Triton X-100 durante 30 min a temperatura ambiente. La incubación se realizó durante toda la noche a 30 °C en tampón acetato sódico 100 mM pH 4.

3.6.4 Actividad coagulante de la leche

3.6.4.1 Actividad de los extractos vegetales

La actividad coagulante de la leche de los extractos vegetales se determinó utilizando una solución al 24 % de leche desnatada en polvo disuelta en tampón fosfato potásico 20 mM, pH 6 y suplementada con una solución CaCl_2 20 mM (Sigma-Aldrich, Steinheim, Germany). La solución se mantuvo en agitación durante 30 min y se preincubó a la temperatura del ensayo. La leche (0.5 mL) fue incubada con los extractos vegetales (0.5 mL) en un baño a 37 °C en tubos de vidrio. Se determinó el tiempo de floculación, definido como el tiempo que transcurre entre la adición del coagulante y la aparición de los primeros flóculos de caseína adheridos a una pared de vidrio (Berridge, 1952)

Los ensayos se realizaron por triplicado; como control negativo se adicionó agua mineral en lugar del extracto vegetal y como control positivo se utilizó quimosina bovina comercial, 0.25 μL Chymax® Plus (Chr. Hansen, Dinamarca) disuelta en 0.5 mL de agua mineral.

Para determinar el mecanismo catalítico de las proteasas presentes en los extractos estos fueron preincubados con diferentes inhibidores de proteasas tal como se indica en el apartado 3.6.1.

3.6.4.2 Actividad de las enzimas recombinantes

La actividad coagulante fue evaluada utilizando el método de Arima et al. (1970) modificado. Leche desnatada en polvo (12 % p/v) fue reconstituida con agua destilada suplementada con CaCl_2 10 mM, y el pH ajustado a pH 6. En un baño de agua a 35 °C se colocaron tubos de vidrio con 1 mL de leche a la que se añadió el coagulante en un volumen final de 50 μL y se determinó el tiempo de coagulación (Berridge, 1952). Todos los ensayos se realizaron por triplicado, utilizando como control negativo 50 μL de agua o tampón y/o el

coagulante preincubado con pepstatina A 1.5 μ M durante 30 min a 37 °C.

La actividad coagulante de la leche (MCA) se expresó en unidades Soxhlet (US) (Harboe et al. 2010), definidas como el volumen de leche coagulada mediante un volumen de coagulante en 40 minutos a 35 °C, mediante la fórmula:

$$US = 2400 V / t * E$$

V representa el volumen de leche (en mL), 2400 representa la conversión de 40 minutos a segundos, t el tiempo de coagulación (en s), y E el volumen de enzima (en mL).

La MCA es expresada como una relación entre el volumen de leche coagulada y el volumen de coagulante utilizado (por ejemplo, 1:5000 implica que 1mL de coagulante es capaz de coagular 5000 mL de leche en 40 min bajo las condiciones del ensayo).

Para comprobar la actividad coagulante de los sobrenadantes sin concentrar obtenidos desde las cepas de *P. pastoris* recombinantes se empleó el método descrito en el apartado anterior (3.6.4.1), debido a la necesidad de emplear un elevado volumen de este en el ensayo para que la coagulación se produjese en un tiempo razonable. El sobrenadante de una cepa de *P. pastoris* transformada con el gen de la quimosina de búfalo (Vallejo et al. 2008) cultivada en las mismas condiciones fue utilizado como control positivo.

3.6.5 Ensayo con BSA como sustrato.

La mezcla de reacción contenía 20 μ L de muestra en una solución tampón citrato sódico 100 mM, pH 5.5 y NaCl 125 mM conteniendo 0.25 μ g de BSA en un volumen final de 80 μ L. La incubación se llevó a cabo a 37 °C durante 2 horas.

3.6.6 Ensayo con FTC-k-caseína como sustrato

El ensayo se realizó mediante el método descrito por (Twining 1984) y modificado por Ageitos et al. (2006). La mezcla de reacción contenía 20 μ L de tampón acetato sódico 100 mM , pH 5.5, 20 μ L de

FTC- κ -caseína y 10 μ L de la muestra enzimática. La mezcla de reacción se efectuó en viales de 1.5 mL que se incubaron en oscuridad a 37 °C y en agitación suave durante 60 a 120 min, dependiendo del ensayo. La reacción se detuvo añadiendo 120 μ L de TCA 5 % frío. Los viales se mantuvieron a temperatura ambiente durante 1 hora, transcurrida la cual se realizó una centrifugación durante 5 min a 18000 g para sedimentar las proteínas insolubles en TCA. Una alícuota de 50 μ L fue diluida en 300 μ L de Tris HCl 0.5 M pH 8.5 y emplazada en una placa de poliestireno Nunc™ F96 MicroWell™ blanca (Roskilde, Denmark). La fluorescencia se midió en un fluorímetro (Luminescence Spectrometer LS50B, Perkin Elmer, Wellesley, MA) con una longitud de onda de excitación y emisión de 490 y 425 respectivamente. Como blanco se utilizó tampón de reacción en lugar de la solución enzimática, la muestra preincubada con pepstatina (1.5 μ M) y/o la muestra añadida a la reacción inmediatamente después de añadir el TCA. Todos los ensayos se realizaron por triplicado.

Una unidad de actividad proteolítica se definió como la cantidad de enzima que genera 100 unidades de fluorescencia transcurridas dos horas de incubación a 37°C.

3.6.7 Ensayo con azocaseína como sustrato

La determinación de la actividad utilizando azocaseína como sustrato se realizó de acuerdo con el protocolo descrito por Akpinar and Penner (2002) con algunas modificaciones. A 20 μ L de enzima se añadieron 480 μ L de una solución azocaseína 1 % (p/v) en tampón fosfato sódico 100 mM pH 6.0 y se incubaron a 37 °C durante dos horas. Cuando se realizó la determinación del pH óptimo los tampones de reacción utilizados fueron: acetato sódico 100 mM pH 4.5 o 5.5, fosfato sódico 100 mM pH 6.0, 6.5 o 7.0. La reacción se detuvo mediante adición de 500 μ L de ácido tricloroacético 10 % (p/v) e incubación durante 20 min a temperatura ambiente. Se realizó una

centrifugación a 8000 g durante 5 min a temperatura ambiente y el sobrenadante fue neutralizado mediante la adición de un volumen NaOH 1 N. La absorbancia de la solución fue determinada a 440 nm utilizando un espectrofotómetro Beckman Du-40. Como blanco se utilizó agua en lugar de la solución enzimática o la mezcla de reacción a la que se añadió la solución enzimática inmediatamente después de la adición del ácido tricloroacético. Todas las determinaciones se realizaron por triplicado. Una unidad de actividad enzimática se definió como la cantidad de enzima que conduce a un incremento de 0.01 unidades de absorbancia en las condiciones del ensayo.

3.6.8 Ensayo con α -, β - y κ -caseínas bovinas como sustrato

α -, β - y κ -caseínas comerciales bovinas (Sigma-Aldrich, St. Louis, MO) fueron disueltas en 100 μ L de tampón fosfato sódico 100 mM pH 6, a una concentración final de 1 mg/mL e incubadas independientemente con 20 μ L de progaine B purificada (1 μ g) o con diferentes coagulantes de la leche comerciales, a 35 °C durante 1 o 2 horas. La reacción se detuvo mediante la adición de tampón de carga del SDS-PAGE (conteniendo β -ME) e incubación a 95 °C durante 7 min. Los productos de digestión se analizaron mediante SDS-PAGE según el método descrito en el apartado 3.4.1.

La cantidad de enzima coagulante añadida a la reacción se calculó aplicando la fórmula indicada en el apartado 3.6.4.2; el cálculo se realizó de forma que todas las muestras analizadas tuviesen la misma fuerza coagulante.

Los coagulantes comerciales utilizados fueron Mucorzyme “L” (Biostar S.A.), quimosina pura (Sigma R4877), Chymax® Plus, Chymax® M (CHR Hansen) y cuajo estándar de ternera y vaca (Calf and Bovine REMCAT Standards; CHR Hansen). Los controles negativos incluyeron el uso de agua en lugar de enzima y el uso de progaine B preincubada con pepstatina A 2 μ M durante 30 min a 37 °C.

Las caseínas hidrolizadas fueron analizadas en un gel SDS-PAGE 16.5 % utilizando tampón de electroforesis Tris-tricina (Tris 100 mM, tricina 100 mM, SDS 0.1 %).

3.6.9 Caracterización de la progaline B

3.6.9.1 pH de activación

El pH de activación de la progaline B fue determinado utilizando el sobrenadante de la cepa recombinante cultivada en YPD7 durante 72 horas. Alícuotas de este sobrenadante fueron mezcladas con un volumen de tampones a diferentes pHs (tampón citrato 0.2 M pH 3; acetato sódico 0.2 M pH 3.6, 4, 4.4 y 5; y fosfato sódico 0.2 M pH 6.5 y 7.0) para obtener soluciones con valores de pH finales de 3.5, 4.3, 4.6, 5.0, 5.8, 6.6, y 7.0 respectivamente. Las muestras fueron incubadas a 28 °C y analizadas transcurridos 30, 150 y 270 min. Alícuotas de 26 µL fueron cargadas en geles de poliacrilamida y analizadas mediante Western blot. Una de las muestras fue tratada con pepstatina A 2 µM a 37 °C durante 30 minutos antes de ser incubada con tampón a pH 3.5. Las muestras incubadas a los diferentes pHs durante 270 min fueron analizadas para determinar su actividad mediante el ensayo FTC-κ-caseína; el pH de las muestras fue ajustado a 5.5 mediante la adición de NaOH o HCl antes de realizar el ensayo.

3.6.9.2 Temperatura óptima y termoestabilidad

El efecto de la temperatura sobre la actividad de la progaline B purificada se determinó mediante el ensayo FTC-κ-caseína. El tiempo de incubación con el sustrato fue de 1 hora:

- La temperatura óptima fue analizada mediante la incubación de la mezcla de reacción a temperaturas que oscilaron entre 30 °C y 80 °C. La actividad se expresó como un porcentaje de la muestra que presentaba una mayor actividad enzimática (actividad relativa).

- La estabilidad de la progaline B al tratamiento térmico fue evaluada mediante su preincubación durante 10 ó 30 min a

temperaturas comprendidas entre 45 °C y 90 °C antes de añadir el sustrato. La actividad proteolítica residual representa la actividad enzimática después del tratamiento con relación a la muestra enzimática no tratada.

3.6.9.3 pH óptimo

La determinación del pH óptimo para la actividad de la progaline B (9 µg) se llevó a cabo utilizando azocaseína como sustrato (apartado 3.6.7), en tampones con valores de pH comprendidos entre 4.5 y 7.5. La actividad se expresó como un porcentaje de la muestra que presentaba una mayor actividad enzimática (actividad relativa).

3.7 Elaboración y análisis de quesos producidos con la progaline B

Para conocer la aptitud de la progaline B para ser utilizada en la industria láctea se elaboraron quesos de pequeño tamaño (~500 g) utilizando leche de vaca cruda o leche de oveja pasteurizada.

El coagulante para la elaboración se obtuvo concentrando el sobrenadante del medio de cultivo (YPD) de la cepa de *P. pastoris* productora de progaline B. Para ello el medio de cultivo de la cepa incubada durante 3 días a 28 °C y 230 rpm fue centrifugado y concentrado (69X) mediante una unidad de ultrafiltración (Model 2000 High Performance Ultrafiltration Cell, Millipore Amicon) utilizando una membrana de límite de peso molecular nominal de 10 kDa (Ultracel de celulosa regenerada; Millipore, Billerica, USA) y filtrado a través de una membrana de 0.22 µm.

El sobrenadante de la cepa silvestre de *P. pastoris* (X-33) cultivada en las mismas condiciones que la cepa recombinante fue concentrado siguiendo el mismo procedimiento que el de la cepa recombinante para ser utilizado como control.

3.7.1 Elaboración de queso con leche de vaca

Se elaboraron dos quesos tipo fresco utilizando el mismo procedimiento, uno con el sobrenadante procedente del cultivo de la cepa de *P. pastoris* productora de progaine B, y el otro con quimosina producida por fermentación de *A. niger* (Chymax® Plus, CHR Hansen), procurando idénticas condiciones en ambas elaboraciones. Los quesos fueron elaborados con leche cruda de vaca (3L) de producción ecológica (leche cruda O Alle, Lalín, Galicia), suplementada con CaCl_2 a una concentración final 5 mM. La leche se calentó en una olla de acero inoxidable, utilizando un baño termostático, hasta alcanzar una temperatura de 35 °C. Una vez alcanzada esta temperatura y mientras se agitaba la leche lentamente se añadió el CaCl_2 y el coagulante. La concentración de los coagulantes añadidos fue calculada para que la coagulación se produjese aproximadamente en una hora, así se utilizaron 100 mL del sobrenadante conteniendo la progaine B o 150 μL de Chymax® Plus disueltos en 40 mL de agua mineral. Una vez formada la cuajada se cortó en cubos de aproximadamente 2.5 cm, para producir el desuerado, y se mantuvo en reposo durante 10 min, tras lo cual se agitó ocasionalmente durante los 15 min posteriores, a los que siguieron otros 15 min de reposo. A continuación, se eliminó parcialmente el suero y la cuajada se introdujo en un colador recubierto con una tela de quesería humedecida y caliente, se permitió el drenaje durante 5 min, se reunieron los extremos de la tela conteniendo la cuajada y se introdujo en un molde de quesería redondeado con tapa de presión. Se realizó un prensado con 3.3 kg de peso durante 3 horas, se volteó el queso y se mantuvo otras 3 horas con el mismo peso. Tras el prensado los quesos se desmoldaron y se sumergieron en una solución de salmuera (20-25 %) durante 5 horas. Los quesos se almacenaron a 4 °C y transcurrida una semana se analizaron.

3.7.2 Elaboración de queso con leche de oveja

Se elaboró un queso tipo curado utilizando leche de oveja de pasto verde (Quesería O Balido, Galicia) y como coagulante el sobrenadante de la cepa de *P.pastoris* productora de progaline B.

La leche se calentó a 63 °C durante 20 min para su pasteurización. Para la elaboración del queso se emplearon 2.5 L de leche a una temperatura de 30 °C suplementada con un cultivo iniciador mesófilo (Choozit TM MA4001 LYO 5 DCU, Danisco), compuesto por las siguientes bacterias: *Lactococcus lactis* subsp. *lactis*, *Lactococcus lactis* subsp. *cremoris*, *Lactococcus lactis* subsp. *lactis* biovar. *Diacetylactis* y *Streptococcus thermophilus*. Transcurridos 30 min la temperatura se subió a 35 °C, se añadió CaCl₂ a una concentración final 10 mM y el coagulante (115 mL). Tras la coagulación se procedió al corte de la cuajada en cubos de 1.5 cm aproximadamente y se dejó reposar hasta que el suero recubrió totalmente la cuajada. Se realizó un segundo corte hasta obtener granos de pequeño tamaño y se agitó ocasionalmente la cuajada durante 15 min, transcurridos los cuales se dejó reposar durante 15 min para favorecer la separación del suero. A continuación, se eliminó el suero utilizando un colador y se permitió el drenaje durante 5 min, se añadieron 6 gramos de sal, se reunieron los extremos de la tela conteniendo la cuajada y se introdujo en el molde de quesería. El prensado se realizó durante 2 horas utilizando un peso de 2 kg, el queso fue volteado y se continuó el proceso de prensado utilizando 4 kg de peso durante 12 horas. La maduración del queso se realizó en las instalaciones de la quesería O Balido, a una temperatura entre los 12 y 14 °C y una humedad del 95 % durante 50 días.

3.7.3 Análisis de las masas queseras

Se realizó un análisis de perfil de textura (TPA, *Texture Profile Analysis*) utilizando un texturómetro TA-XT2i Stable Micro Systems (Vienna Court, England). El análisis se llevó a cabo

utilizando una sonda cilíndrica de aluminio de 50 mm de diámetro (P/50). Las muestras se sometieron a un porcentaje de compresión del 75 % con respecto a la altura inicial, utilizando una velocidad de compresión de 60 mm/min. Mediante esta prueba se determinaron los parámetros de dureza, adhesividad, elasticidad, cohesividad, gomosidad y masticabilidad.

Las cantidades de proteína (AOAC 981.10), cenizas (AOAC 920.153) y grasas (AOAC 991.36) fueron obtenidas siguiendo los procedimientos de la “Association of Official Analytical Chemists” (AOAC 2002). Los hidratos de carbono y el contenido energético fueron determinados según el procedimiento recomendado por Allison y Senti (1983); los ácidos grasos fueron determinados según la metodología establecida por Martínez y colaboradores (2012).

Los análisis se realizaron en el Departamento de Química Analítica, Nutrición y Bromatología de la Universidad de Santiago de Compostela.







RESULTADOS



4. RESULTADOS

4.1 Estudios de la actividad coagulante de la leche

Se analizó la actividad coagulante de la leche de las especies vegetales objeto de estudio, y el tipo de proteasas que estaban implicadas en la coagulación.

4.1.1 Análisis de la actividad coagulante de *G. verum*

Se comprobó que la planta *G. verum* sometida a un proceso de secado tenía actividad coagulante de la leche, y que esta actividad se detectaba en los extractos realizados con las flores y hojas, mientras que estaba ausente en los tallos (Figura 17). El tiempo de coagulación fue de 2 horas y 15 minutos cuando se utilizaron los extractos procedentes de las hojas, mientras que utilizando extractos obtenidos a partir de la planta completa o de las flores fue ligeramente superior (2 horas 30 minutos). El pH del extracto elaborado con las flores fue de 4.8, mientras que el procedente de las hojas era de 4.5, lo cual podría influir en el tiempo de coagulación. Cuando los extractos eran preincubados con el inhibidor de proteasas aspárticas pepstatina A, la leche permanecía líquida después de 6 horas de incubación (Figura 17, tubos 4 y 5), mientras que cuando los inhibidores utilizados eran de serín- o cisteín-proteasas, el tiempo transcurrido hasta que se producía la coagulación era el mismo que el de los extractos a los que no se les había añadido ningún inhibidor (Figura 17 tubos 6 y 7). Estos resultados indican, como era de esperar, que *G. verum* tiene actividad coagulante de la leche y que esta es debida a proteasas de tipo aspártico.

La actividad coagulante de la preparación cruda fue menor que la obtenida con los extractos secos. Así, la coagulación no se producía hasta que habían transcurrido 5 horas. Al igual que ocurría con los

extractos secos, la leche no coagulaba cuando se realizaba la preincubación del extracto con pepstatina.

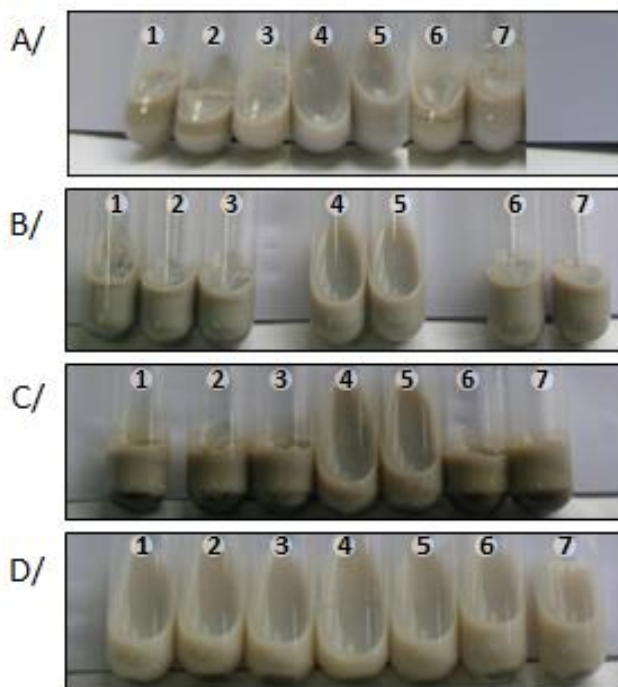


Figura 17. Análisis de la capacidad coagulante de la leche de extractos vegetales de la planta *G. verum* desecada: (A) planta completa, (B) flores, (C) hojas y (D) tallos. La actividad se ensayó incubando los extractos (tubos 1-3) con leche descremada suplementada con calcio. El efecto de diferentes inhibidores se analizó preincubando los extractos con pepstatina A (tubos 4 y 5), PMSF (tubo 6) y E-64 (tubo 7) antes de ser añadidos a la leche.

4.1.2 Análisis de la actividad coagulante de *F. carica*

El látex procedente de la especie vegetal *F. carica*, clarificado mediante centrifugación, producía la coagulación instantánea de la leche. La coagulación se producía aun cuando el látex era preincubado en presencia de inhibidores de diferentes tipos de proteasas; cuando se calentaba a 100 °C la leche permanecía líquida. La elevada actividad proteolítica del látex hizo necesario realizar una dilución 1/1000 del

extracto (concentración proteica final 1.5 mg/mL) para poder evaluar el grupo mecanístico al que pertenecían las proteasas presentes en el mismo. El extracto diluido producía la coagulación de la leche transcurrido 1 min de incubación a 37 °C; cuando el extracto era preincubado con el inhibidor de proteasas aspárticas pepstatina A, el tiempo de coagulación apenas se veía afectado. La incubación con el inhibidor de serín proteasas, PMSF, provocaba que la coagulación no se produjese hasta que habían transcurrido 9 min, mientras que cuando se añadía el inhibidor de cisteín-proteasas, E-64, la combinación de este con PMSF o una combinación de los tres inhibidores testados, la coagulación no se producía hasta que habían transcurrido seis horas. Este ensayo indica que la mayoría de las proteasas con actividad coagulante de la leche presentes en el látex pertenecen probablemente al grupo de las cisteín-proteasas.

4.2 Búsqueda de genes codificantes de proteasas aspárticas de *G. verum* y *F. carica*

Una vez comprobada la actividad coagulante de la leche de las especies vegetales objeto de estudio, abordamos la búsqueda de genes codificantes de PAs presentes en su genoma.

4.2.1 Obtención de secuencias génicas

4.2.1.1 Secuencias de los genes *gv1AP*, *gv2AP* y *fc1AP*

La identificación de las secuencias de los genes codificantes de proteasas aspárticas se realizó en dos pasos:

- (i) Primero se obtuvieron secuencias parciales mediante una RT-PCR empleando dos cebadores degenerados, diseñados en base a dos secuencias conservadas de PAs vegetales depositadas en la base de datos del NCBI.
- (ii) Segundo, la secuencia completa de los genes se obtuvo utilizando la técnica “Rapid amplification of cDNA ends” (RACE).

Los cebadores degenerados (pAPF y pAPR) se diseñaron en base a dos secuencias de aminoácidos (DTGSSNLW y WILGDVFM) conservadas en varias proteasas aspárticas vegetales clasificadas como típicas (Figura 18).



Figura 18. Alineamiento utilizado para el diseño de los cebadores degenerados pAPF y pAPR. En el margen izquierdo se indican los códigos Genbank de las PAS vegetales: *S. lycopersicum* (AAB18280), *C. vulgare* (AFB73927), *S. marianum* (AGE15494), *C. cardunculus* (CAA48939) y *C. calcitrapa* (CAA70340). Los residuos conservados se encuentran sobre fondo azul y los residuos idénticos sobre fondo amarillo. Los recuadros enmarcan las secuencias utilizadas para el diseño de los cebadores. Los números indican las posiciones de los aminoácidos en el alineamiento múltiple.

La integridad del ARN total extraído desde la planta florida de *G. verum* y de los siconos inmaduros de *F. carica* se analizó mediante electroforesis en geles de agarosa en condiciones desnaturalizantes. En la Figura 19 se observan claramente las bandas correspondientes a las subunidades intactas del ARN ribosómico de *G. verum*.

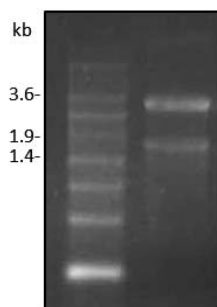


Figura 19. Electroforesis en gel de agarosa de ARN. Izq.: marcador de pesos moleculares de ARN, dcha.: ARN total extraído desde *G. verum*.

El ARN extraído fue utilizado para la síntesis de la primera hebra de ADNc empleando como cebadores oligo(dT)₁₅, hexanucleótidos pd(N)₆ de secuencia al azar o el cebador degenerado pAPR. El ADNc sintetizado fue empleado para realizar una PCR con los cebadores degenerados pAPF y pAPR. En la Figura 20 puede apreciarse como esta estrategia dio lugar a la amplificación de fragmentos de un tamaño de 1.1, 1, 0.7 y 0.4 kb en el caso de *G. verum* y de 1.2, 0.9, 0.7, 0.6 y 0.4 kb en el caso de *F. carica*. Los fragmentos de PCR con un tamaño de 1.1 kb en el caso de *G. verum* y de 1.2 kb en el caso de *F. carica* fueron purificados a partir del gel de agarosa, el resto de las bandas que aparecen en ambos geles se descartaron debido a que (i) no tenían el tamaño del amplicón esperado y que (ii) las bandas del tamaño seleccionado estaban presentes en todas las PCRs realizadas independientemente del cebador utilizado para la síntesis del ADNc o de la temperatura de hibridación empleada (Figura 20).

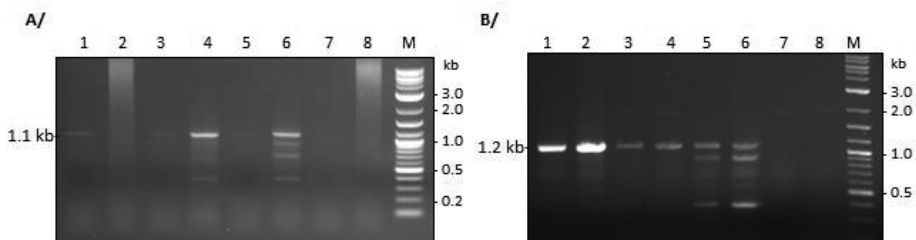


Figura 20. Amplificación de secuencias internas que codifican para proteasas aspárticas de *G. verum* (A) y *F. carica* (B). Los productos de la RT-PCR fueron obtenidos utilizando los cebadores oligo(dT)₁₅ (calles 1 y 2), hexanucleótidos al azar (calles 3 y 4) o pAPR (calles 5 y 6). Las temperaturas de hibridación fueron 55 °C (calles 1, 3 y 5) y 49 °C (calles 2, 4 y 6) en el caso del ADNc de *G. verum* (A) y 50 °C (calles 1, 3 y 5) y 47.9 °C (calles 2, 4 y 6) en el caso de *F. carica* (B). En las calles 7 y 8 de ambos geles (A y B) se utilizó como molde ARN al que no se le añadió transcriptasa inversa, para verificar que el producto amplificado no provenía de ADN genómico. M indica el marcador de pesos moleculares. Los números en los márgenes de ambos geles representan los pesos moleculares del marcador (dcha.) y de los fragmentos seleccionados para la purificación (margen izq.)

La clonación de los fragmentos obtenidos en el vector pCR[®]-Blunt II-TOPO[®] nos permitió obtener dos secuencias parciales que codificaban para posibles proteasas aspárticas de *G. verum* y tres secuencias parciales para la especie *F. carica*.

A partir de estas secuencias se diseñaron cebadores específicos (Tabla 10) que nos permitieron, utilizando la técnica RACE PCR (Figura 21) (descrita en materiales y métodos, apartado 3.2.10.2), la obtención de los extremos 5' y 3' de la región codificante de dos genes, *gv1AP* y *gv2AP*, de *G. verum* y de uno, *fc1AP*, de *F. carica*, así como la secuencia parcial de dos genes de esta última especie, que denominamos *fc2AP* y *fc3AP*.

La secuencia de los extremos 5' y 3' del gen *fc2AP* no ha podido ser determinada a pesar de haber utilizado el mismo protocolo que el descrito para los otros genes; la obtención de la secuencia del gen *fc3AP* no se ha abordado en este estudio.

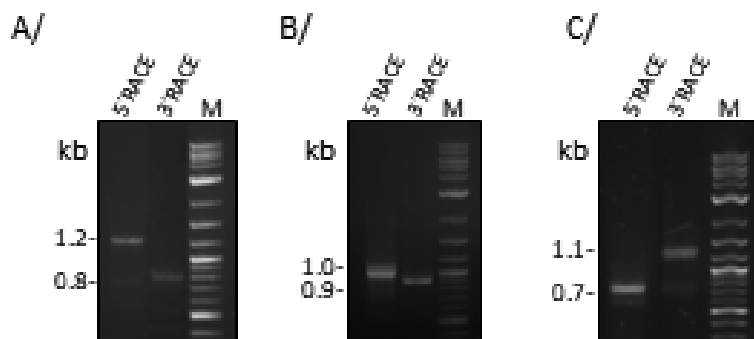


Figura 21. RACE PCR realizada a partir de ADNc de *G. verum* y *F. carica*. (A) PCR a partir del ADNc de *G. verum* para la obtención del gen *gv1AP*: cebadores AP1/pMgv1r (5' RACE) y pMgv1f/AP1 (3' RACE), (B) del gen *gv2AP*: cebadores AP1/pMgv2r (5' RACE) y pMgv2f/AP1 (3' RACE) y (C) del gen *fc1AP* a partir de ADNc de *F. carica* con los cebadores AP1/pMfc1r (5' RACE) , pMfc1f/AP1 (3' RACE).

Las secuencias de nucleótidos obtenidas han sido incorporadas a la base de datos pública Genbank, con los nombres y códigos de acceso que figuran en la Tabla 11.

Tabla 11. Códigos de acceso de las secuencias de nucleótidos de *G. verum* y *F. carica* depositadas en el GenBank.

Nombre	Código Genbank	Longitud (nt)
<i>Galium verum</i> preprogaline A (<i>gv1AP</i>) mRNA, complete CDS	JQ434478	1,524 pb
<i>Galium verum</i> preprogaline B (<i>gv2AP</i>) mRNA, complete CDS	JQ434479	1,722 pb
<i>Ficus carica</i> preprofigain A (<i>fc1AP</i>) mRNA, complete CDS	JQ434480	1,506 pb
<i>Ficus carica</i> preprofigain B (<i>fc2AP</i>) mRNA, partial CDS	JQ434481	1,155 pb
<i>Ficus carica</i> preprofigain C (<i>fc3AP</i>) mRNA, partial CDS	JQ434482	1,158 pb

La técnica RACE PCR nos permitió obtener parte de las secuencias no codificantes de los extremos del ARNm, UTR's (*UnTranslated Region*) de los genes *gv1AP*, *gv2AP* y *fc1AP* (Tabla 12). Las regiones 5' y 3' no traducidas (UTRs) de los ARNm eucarióticos desempeñan papeles cruciales en la regulación postranscripcional de la expresión génica a través de la modulación del transporte de ARNm núcleo-citoplasmático, la eficacia de la traducción, la localización celular o la estabilidad del ARNm (Mignone et al. 2002). Los UTRs obtenidos probablemente estén constituidos por un mayor número de nucleótidos de los que figuran en la tabla, ya que la longitud media de estas secuencias en plantas se encuentra entre 100-200 pb para el extremo 5'UTR y 200-500 pb para el extremo 3'UTR (Mignone et al. 2002; Srivastava et al. 2017). Al igual que ocurre en el 80 % de los genes de plantas, los nucleótidos que flanquean al codón de inicio AUG (subrayado y resaltado en la Tabla 12), son una purina en posición -3 y una guanina en posición +4 , con excepción del gen *fc1AP* que presenta en posición +4 el nucleótido uracilo (Joshi et al. 1997). Ninguno de los genes presenta el nucleótido A/C en posición -2, que podría estar implicado según un estudio *in silico* en la eficiencia de la traducción en genes de origen vegetal (Gupta et al. 2016).

Tabla 12. Secuencias UTR's obtenidas desde el ADNc utilizando la técnica RACE PCR.

Gen	UTR	Secuencia
<i>gv1AP</i>	5'UTR	AAAUCCCACCAAGGUUGUGUGAGU <u>AUG</u> GAG
	3'UTR	GCCUAUGGUGUUUCGAGUAUUACCGAACUCGAAUUGUUUGAU GGAAUGAUCACUCAAGUUAAGAUGAAAAACAGUGUGAGGGUA AUAUAUCUUGCUUUGGCUUUCUUUUGGCUAAAGCAUAAAAUGUA AUAGUAUGAUGAGAUGAUCAUCCUGUAUUUGUAUCUUGAUG UACUGAUUUUCAGAUAGAUUAUGUAUUUAUUACACUAAAAU ACUUAUACACUUAUAACUUAUAUGGUUAUUCUACUUUGUUC U
<i>gv2AP</i>	5'UTR	CCAGCUCUUUGCAAUUGGG <u>AUG</u> GGCUC
	3'UTR	GAAUACUCAGCAUAUCAAUAAUGGAUGUUUAAUUAACCGCG GUCAUUAUGCAGUUGUUCUCUGCACCACUACCCUCUGUUUGU UAUACUGUACAUAUAUUGUUCGGUUUCACUCCAACUGUUAC UGCUUCGUGUUUAAUUUAAAGCGACUAAUUUCAGUCCAUUU CUU
<i>fc1AP</i>	5'UTR	UCCCATAUUAAAAAAAAAAAAAAAAAAAAAACUCUUUUGAGCU UCAUCUCCACUCUCUCAACACAAAGGUUUUUUUCUAAACAU <u>G</u> <u>AUG</u> UGGCA
	3'UTR	AGGCGUUUGCCGCUUCUUUAUCUGUUGUGUGUCUAAAUUU UAGAGACCUUUUUUUUAUAGUAGUUUGGGAUGGUUAUAG AAAAAACAAUGCAGGAGACGUUACUACCUAAAUGCUGGUUGCC AAAUUACAUUUUGAGAUCUUAUGAAGAUCAUUAUGUCAUGU

4.2.1.2 Análisis parcial de los intrones del gen *gv1AP*

Se abordó el estudio de la secuencia del genoma de *G. verum* que codifica para el gen *gv1AP* con el fin de determinar su estructura. Se realizó una PCR utilizando el ADN genómico y las parejas de cebadores pGv1-EcoRI/pMgv1r y pMgv1r /pGv1-XbaI (Tabla 10) que abarcan la región codificante del gen. Las secuencias obtenidas se

compararon con las secuencias de ADNc obtenidas previamente para este gen y se analizaron con ayuda del programa informático NetGene2 Server, determinándose así, sin ambigüedades, las uniones exón-intrón. No se obtuvo la secuencia genómica completa de la región codificante, únicamente la que abarca la región del CDS comprendida entre el nucleótido 824 (secuencia JQ434478) y el codón de terminación. Esta región se encuentra interrumpida por seis intrones, todos ellos presentan la secuencia GU (Guanina-Uracilo) en el extremo 5' y la secuencia AG (Adenina-Guanina) en el extremo 3' (Brown y Simpson 1998), cumpliendo la regla GT-AG de la mayoría de los genes nucleares de eucariotas (Figura 22). Se comparó la disposición y tamaño de exones-intrones del gen *gvIAP* con la que presentan los genes que codifican para tres PAs típicas, el gen de la cardosina B de *C. cardunculus* (Genbank: AM286227; Pimentel et al. 2007), gFeAP9 de *F. esculentum* (Genbank: AM422870; Timotijević et al. 2010) y oryzasin de *O. sativa* (Genbank: D32165; Asakura et al. 1995), comprobándose que la interrupción de los exones ocurría en posiciones muy conservadas (Figura 23.A), tanto es así que, el aminoácido triptófano (W), situado entre los exones que denominamos E_E/E_F en el caso de *gvIAP* o E12/E13 en el resto de las proteínas alineadas, estaba codificado por el codón TGG, perteneciendo la timina a un exón y las dos guaninas al posterior en todas las proteasas analizadas. El gen *gvIAP* carecía del intrón situado entre los exones E11 y E12, esta región, correspondiente al final del PSI y a la región inicial de la cadena ligera, estaba codificada por un único exón (E_E), el resto de los exones ocupaban posiciones homólogas a los genes analizados (Figura 23.A). El tamaño de los intrones variaba entre las diferentes especies (Figura 23.B).

1 aaacttcttc ncaagtagtt atctattagtt ttttctggtt gttatctgccc ataattttat attatatcaa aaataattcaa ggacaaacat cctcactgct
101 tctaaatac tctgatacgt ctcataatc taagctcggg tttggatatt gttgtatttc tccaataggt TTCGTGAAG GAGGTGTAA GGCTATTGTG **EA**
201 GACTCTGGAA CATCATGGT TGTGGCCCA ACCgtatggt cagataaact tcaattttat gacaatttcaa taaatttgaa atcattgagct atatttgctc
301 ttgcaaaaca taaacttcgt aggcgtcctt ttacagacta TCGTGACTGA AATCAACCAC GCCATTGGTG GAGAAGAGT GGTAAACGGTG GAGTGCARAAG **EB**
401 EAATTGTTTC AGAGTATGGA GACATGATAT GGAATCTTCT TGTATCAGGG gtctgttctt tctctctcc tocattctcg tgggttaatt attattttt
501 tttttttttt tttttggcta aaatattcag gcaagtgatt ctgaattaaa atttttggtt tcataggttt CACCTGACAA GGTCGGGCA CAAGCTGGTT **EC**
601 L C L F A E A Q S D S TATGTTTATT CGCTGAAGT CAGTCTGACA Ggttagcaaa tcccgagttt tgctatctat gacctocaaac tgcctatctg atgcataac ctccaagtta
701 cctcatatgt aattaaaatc acaaaatctg ttcaaaatta tttcgtgaag agctcttttg gttcaggaag ttgagatttt tggttcaagt actttgatat
801 gaactgaag tcttgattat tcttgttcat catacttggt tggdtaatat gctgatatgc ttttgattat ggcagtttaca TCAAGTCGGT TGTGTATGAT **ED**
901 E N K D G H S V G I T P M C S A C Q M A V W I R N Q L K Q K N T GAGGAGAATA AGGACGGCA TTCTGTTGGT ATTACACCGA TGTGTTCTGC TTGCCAGATG GCAGTTGTTT GGATTGGAA CCAGGTGAAA CAGAAGAACA
1001 T K D K V F E Y V N Q CCAAGGACAA AGTGTTTGAA TATGTGAATC agttaaactg taaacaagta tcaattttct ataaaccta aatttagcca agatcagtta coatacatt
1101 L C E S I P S P M G E S V V S C N S L S S L P N I S tatgtttttt caatgtttct cagctctctg AGACATTC GAGCCCATG GGTGAATCTG TAGTCAGCTG TAACAGCTA TCCAGCTTC CAAACATATC **EE**
1201 S F T I G D K V F E L T P E Q Y V L K T G T G M E T I C I S G F M A ATTACCATT GGAGATAAAG TTTTCGAGTT AACCCAGAA CAATATGTTT TTAACCCGG AACCTGGATG GAAACTATTT GCATCAGCGG ATTCAATGGCT
1301 F D L P A P R G P L M TTTGACTTAC CTGCCCCAG CGGTCTCTC Tggtacgcca acctgttca aattcaatcg atgaaattac attctcctc tttttcaac tgctgtcaaa
1401 atggctgat gatacttact ttacgcctt catggattgg tgtattgttg ctgtagat TCTTGGGAC GTATTCATGG GGGTATACCA CACTGTGTTC **EF**
1501 D Y G N L Q L G F A K A * GATTATGGCA ACCTCCAGCT AGGATTGCGG AAGGCTGCTTAG

Figura 22. Secuencia del ADN genómico de *G.verum* que codifica para el gen *gv1AP*. Los nucleótidos pertenecientes a los exones se encuentran subrayados en gris y en mayúsculas, los correspondientes a los intrones se muestran en letras minúsculas. Los sitios de unión exón-intrón siguen la regla GT-AG (subrayados en amarillo). Los seis exones (E_x) se corresponden con los siguientes nucleótidos de la secuencia depositada en el Genbank (JQ434478): E_A (830-894); E_B (895-1008); E_C (1009-1073); E_D (1074-1230); E_E (1231-1439); E_F (1440-1524). Los aminoácidos (en rojo) codificados por los exones se muestran en código de una letra. Los nucleótidos subrayados pertenecen a un codón que se encuentra en dos exones diferentes. El símbolo * indica el codón de terminación.

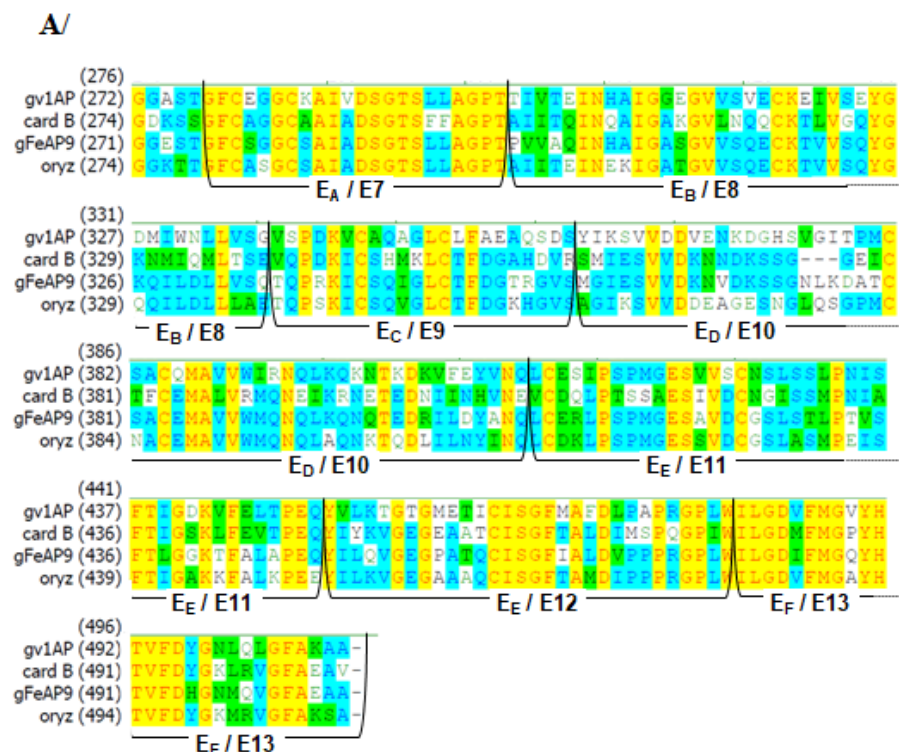
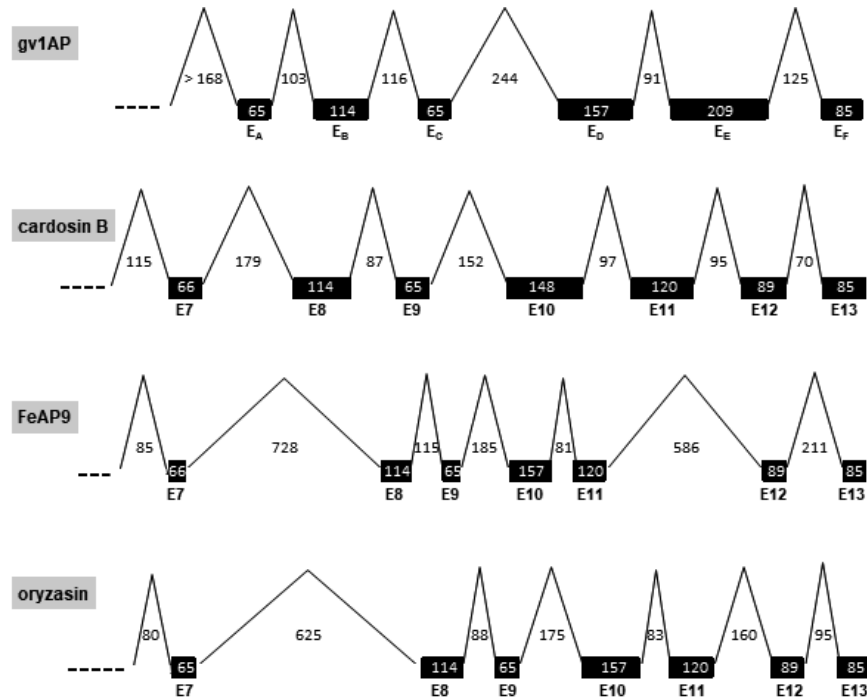


Figura 23. Comparación de la estructura exón-intrón del extremo 3' del gen *gv1AP* con genes de otras PAs vegetales típicas. **(A)** Alineamiento del extremo C-terminal de las proteínas codificadas por los genes *gv1AP* (Genbank: AFX73038), cardosina B (*card B*) (Genbank: CAB40349), *gFeAP9* (Genbank: CAM12780) y *oryzasin* (*oryz*) (Genbank: BAA06875). En azul se encuentran resaltados los residuos conservados, en verde los residuos similares y en amarillos los residuos idénticos. Delimitados por líneas negras se encuentran los residuos pertenecientes a cada exón, abreviados como E y una letra en el caso del gen *gv1AP*, o E y el número de exón en el caso de la cardosina B, *gFeAP9* y *oryzasin*. **(B)** Representación esquemática de los genes: los rectángulos negros representan exones y los triángulos abiertos simbolizan intrones. El tamaño de ambos elementos se indica en pb. El número de exones indicados se refieren a la región codificante, no contabilizando el exón del extremo 5' UTR presente en los genes cardosina B, *oryzasin* y *gFeAP9*; el número de nucleótidos indicados en E_F y E13 de los genes no incluye los pertenecientes a la región no codificante 3' UTR.

(continuación Figura 23)

B/

4.2.1.3 Localización de los tejidos de expresión del gen *fc1AP*

Mediante RT-PCR del ARN mensajero extraído a partir del interior de los siconos, hojas y pedúnculos de *F. carica* comprobamos que el gen *fc1AP* se estaba expresando en el momento de la recolección en todos los tejidos analizados (Figura 24).

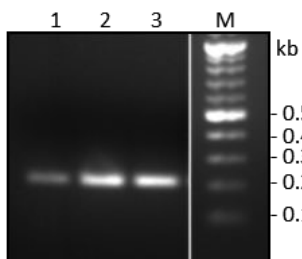


Figura 24. Análisis de la transcripción del gen *fc1AP* en diferentes tejidos. En la imagen se aprecia la banda de 192 pb correspondiente al producto de la RT-PCR del sicono (calle1), hoja (calle2) y pedúnculo del sicono (calle3) de *F. carica* utilizando los cebadores pMfc1f y pMfc1r. Los números en el margen derecho representan los pesos moleculares del marcador.

4.2.2 Análisis bioinformático de las secuencias

El análisis bioinformático de las secuencias identificadas mediante RACE PCR determinó que la secuencia *gv1AP* codifica para una proteína, que denominamos preprogaline A, de 507 aminoácidos, *gv2AP* codifica para una proteína, preprogaline B, de 506 aminoácidos, y en el caso de *fc1AP* la proteína resultante, denominada, preprofigain A, está constituida por 501 aminoácidos. Las propiedades fisicoquímicas obtenidas mediante el programa ProtParam del servidor ExPASy permitieron establecer el peso molecular y el punto isoelectrico de estas proteasas (Tabla 13).

Tabla 13. Predicción de parámetros fisicoquímicos de las proteasas aspárticas identificadas.

Proteasas	Peso molecular (kDa)	Punto isoelectrico
Preprogaline A	55.00	5.38
Preprogaline B	54.39	5.33
Preprofigain A	54.89	5.00

4.2.2.1 Análisis estructural

El alineamiento de las secuencias de aminoácidos deducidas a partir de los genes *gv1AP*, *gv2AP* y *fc1AP*, revela una elevada similitud con las de otras especies vegetales pertenecientes a la familia A1 de la base de datos MEROPS, como la fitepsina de la cebada, la AtAP1 de *A. thaliana*, SoyAP1 de la soja, StAP de la patata o la cardosina A y ciprosina B procedentes del cardo (identidad superior al 54 % en todas ellas). Las mayores diferencias se encontraron en la región N-terminal, correspondiente al péptido señal y al prosegmento, y en la región del PSI. La similitud con la secuencia de la proteasa aspártica de origen animal quimosina, utilizada para la fabricación del queso, es también elevada si omitimos el PSI al realizar el alineamiento (Figura 25).

	Péptido señal	Prosegmento
	→
Fitepsina	(1) -----MGTRGIALALAAVLLLTQVLPAASEAEGIVRIALKKRPIDRNSRVATGLSGGE-----EQPLLSGANPLRSE-----EEGDIVALKNYVNAQYFGEIGV	
AtAP1	(1) MYTYWLTFMGKMKIYSRTVAIVSLIVSFLLCFSFAERNDDGTFRRVGLKKLLDLSKNRLAARVRSKQ-----EKLRAVRLGSDG-----ADVVLKNYLDAQYYGEIAI	
SoyAP1	(1) -----MGNRNVAIVLCLVSTLLVSAVYCAPNAGRRIRIGLKKIKLDPKNRLAARVRSKQ-----VDSFRASIRQPHLQNNFGGTEEDIVALKNYLDAQYYGEIAI	
StAP	(1) -----MDKKHLCAALLLWATCSALPASSGDLRLIGLKKHFLDINSIKAAARVAKLQ-----DRYGGKHVINEKKSS-----DSDIDIVLPLKNYLDAQYYGEIGI	
Cardosin A	(1) -----MGTSIKANYLALFLFYLLSPITVFSVSDGILIRIGLKKRKVRIDQLGRRAALME-----GNARKDFGFRGTVRDSG-----SAVVALTNDTRDTSYFGEIGI	
Ciprosina B	(1) -----MGTAIKASVIALFLFFLLSPITAFSVNGGLLRVGLKKRKVDQINQLSGHGVSMW-----AKARKDFGFGALRDSG-----SDIIALKNYVDAQYYGEIGI	
Preprogaline A	(1) -----MEHRYLLVALFLSAICSLFFPYFSEGVLRIGLKKSLDLQGVRAILARSRPHGNCRKLNSQLSSSS-----GDILPLKNYLDAQYYGDIISI	
Preprogaline B	(1) -----MGSTLRATIVAFLLAFVLSAALLSASETKIVRVSLKKAKLDQVNRLAGSVPRGS-----ARINRLYGNMGSSD-----SDIIALKNYVDAQYYGEIGI	
Preprofigain A	(1) -----MWHKHLLAGFYIWAIVICSLLPDSSNGLMRLNKKRALDLESIKADRIVR-----ESKYKEFHHLGNSD-----EDIVLKNYLDQAQYYGEIGI	
Preprofigain B	(1) -----	
Preprofigain C	(1) -----	
Quimosina	(1) -----MRCLVLLAVATPALSQGAETIRIPYKKGSLRALKHEHGLLEDFLQKQQ-----YGISSKYSYSGFGEVAS-----VPLTNLYLDSQYFCKTYL	

	...	^	^
Fitepsina	(91) GTPPQKFTVIFDTGSSNLWVPSAKCYFSIACVYIHSRYKAGASSTYKKNKCPAAIQYGTGSIACYFSDSVTVGDIIVVKQEFTEATKPPGITEFLVAKFDGELGLGFKEIS		
AtAP1	(100) GTPPQKFTVIFDTGSSNLWVPSAKCYFSIACVYIHSRYKAGASSTYKKNKCPAAIQYGTGSIACYFSDSVTVGDIIVVKQEFTEATKPPGITEFLVAKFDGELGLGFKEIS		
SoyAP1	(97) GTSPQKFAVIFDTGSSNLWVPSAKCYFSIACVYIHSRYKAGASSTYKKNKCPAAIQYGTGSIACYFSDSVTVGDIIVVKQEFTEATKPPGITEFLVAKFDGELGLGFKEIS		
StAP	(90) GSPPKFKVIFDTGSSNLWVPSAKCYFSIACVYIHSRYKAGASSTYKKNKCPAAIQYGTGSIACYFSDSVTVGDIIVVKQEFTEATKPPGITEFLVAKFDGELGLGFKEIS		
Cardosina A	(92) GTPPQKFTVIFDTGSSNLWVPSAKCYFSIACVYIHSRYKAGASSTYKKNKCPAAIQYGTGSIACYFSDSVTVGDIIVVKQEFTEATKPPGITEFLVAKFDGELGLGFKEIS		
Ciprosina B	(92) GSPPKFKVIFDTGSSNLWVPSAKCYFSIACVYIHSRYKAGASSTYKKNKCPAAIQYGTGSIACYFSDSVTVGDIIVVKQEFTEATKPPGITEFLVAKFDGELGLGFKEIS		
Preprogaline A	(90) GSPPKFKVIFDTGSSNLWVPSAKCYFSIACVYIHSRYKAGASSTYKKNKCPAAIQYGTGSIACYFSDSVTVGDIIVVKQEFTEATKPPGITEFLVAKFDGELGLGFKEIS		
Preprogaline B	(89) GTPAQKFTVIFDTGSSNLWVPSAKCYFSIACVYIHSRYKAGASSTYKKNKCPAAIQYGTGSIACYFSDSVTVGDIIVVKQEFTEATKPPGITEFLVAKFDGELGLGFKEIS		
Preprofigain A	(84) GSPVQNFVIFDTGSSNLWVPSAKCYFSIACVYIHSRYKAGASSTYKKNKCPAAIQYGTGSIACYFSDSVTVGDIIVVKQEFTEATKPPGITEFLVAKFDGELGLGFKEIS		
Preprofigain B	(1) -----GSSNFWVPSAKCYFSIACVYIHSRYKAGASSTYKKNKCPAAIQYGTGSIACYFSDSVTVGDIIVVKQEFTEATKPPGITEFLVAKFDGELGLGFKEIS		
Preprofigain C	(1) -----DTGSSNLWVPSAKCYFSIACVYIHSRYKAGASSTYKKNKCPAAIQYGTGSIACYFSDSVTVGDIIVVKQEFTEATKPPGITEFLVAKFDGELGLGFKEIS		
Quimosina	(81) GTPPQKFTVIFDTGSSNLWVPSAKCYFSIACVYIHSRYKAGASSTYKKNKCPAAIQYGTGSIACYFSDSVTVGDIIVVKQEFTEATKPPGITEFLVAKFDGELGLGFKEIS		

Figura 25. Alineamiento de la secuencia de aminoácidos de los genes obtenidos en este estudio con otras PAS vegetales. Los aminoácidos correspondientes al péptido señal se encuentran bajo una flecha, el prosegmento bajo una línea de puntos y los pertenecientes al PSI en un recuadro. Los motivos catalíticos aparecen indicados sobre la primera secuencia del alineamiento con puntos negros, los residuos de cisteínas conservados aparecen bajo el símbolo ^ y los residuos Lys/Tyr bajo el símbolo *. Los guiones insertos en las secuencias indican posiciones en las que no existen residuos homólogos. Los residuos conservados en todas las secuencias se encuentran sobre fondo negro y los residuos idénticos sobre fondo gris. Los números indican las posiciones de los aminoácidos en el alineamiento. Los códigos de acceso de las secuencias fueron: fitepsina (CAA39602.1); AtAP1 (ANM60511.1); SoyAP1 (BAB62890.1); StAP (AAT77954.1); cardosina A y ciprosina (CAB40134 y CAA57510); preprogaline A y preprogaline B (AFX73038.1 y AFX73039.1); preprofigain A, B y C (AFX73040.0, AFX73041.1 y AFX73042.1); se incluye la secuencia de la quimosina (Bos taurus, ABU41411). El alineamiento se realizó con el programa VectorNTI.

Fitepsina	(201)	VGVKAVFVWYKMI	EQGLVSDPVSF	FWLNRHVDGEGEII	IFGGMDPKHYVE	HTYVPV	OKKQWQF	DMGDVL	VGGKSTGF	CGAGGAIAI	DSGSL	LAGPT	AIITE	INFKIG
AtAP1	(210)	VGVKAPFWYNM	LKQGL	IKEPV	FSFWLNRHVDGEGEL	VFGV	DPNHF	KCHTYVPV	OKKQWQF	DMGDVL	IGAP	TGFC	SGGSAI	DSGSL
SoyAP1	(207)	VGNVAPFWYNM	VQGL	LKEPV	FSFWLNRHVDGEGEII	VFGV	DPNHF	KCHTYVPV	OKKQWQF	DMGDVL	IGAP	TGFC	SGGSAI	DSGSL
StAP	(188)	VGNVTPFWYNM	VQGL	LKEPV	FSFWLNRHVDGEGEL	VFGV	DPNHF	KCHTYVPV	OKKQWQF	DMGDVL	IGAP	TGFC	SGGSAI	DSGSL
Cardosina A	(202)	---	VF	WYNM	VQGL	LKEPV	FSFWLNRHVDGEGEL	VFGV	DPNHF	KCHTYVPV	OKKQWQF	DMGDVL	IGAP	AIITE
Ciprosina B	(202)	VGVSVBELWYNM	VQGL	LKEPV	FSFWLNRHVDGEGEL	VFGV	DPNHF	KCHTYVPV	OKKQWQF	DMGDVL	IGAP	TGFC	SGGSAI	DSGSL
Preprogaline A	(200)	VDDMWELWYNM	VQGL	LKEPV	FSFWLNRHVDGEGEL	VFGV	DPNHF	KCHTYVPV	OKKQWQF	DMGDVL	IGAP	TGFC	SGGSAI	DSGSL
Preprogaline B	(199)	VGVKAPFWYNM	VQGL	VAEPV	FSFWLNRHVDGEGEL	VFGV	DPNHF	KCHTYVPV	OKKQWQF	DMGDVL	IGAP	TGFC	SGGSAI	DSGSL
Preprofigain A	(194)	VGNVTPFWYNM	VQGL	VAEPV	FSFWLNRHVDGEGEL	VFGV	DPNHF	KCHTYVPV	OKKQWQF	DMGDVL	IGAP	TGFC	SGGSAI	DSGSL
Preprofigain B	(99)	VGNVAPFWYNM	VQGL	LKEPV	FSFWLNRHVDGEGEL	VFGV	DPNHF	KCHTYVPV	OKKQWQF	DMGDVL	IGAP	TGFC	SGGSAI	DSGSL
Preprofigain C	(100)	VGNVAPFWYNM	VQGL	LKEPV	FSFWLNRHVDGEGEL	VFGV	DPNHF	KCHTYVPV	OKKQWQF	DMGDVL	IGAP	TGFC	SGGSAI	DSGSL
Quimosina	(190)	SEYSIE	VF	DMNRH	VAQDI	ESV	YVMDR	NGQES	--	MLTICA	IPSY	YV	YV	YV
Fitepsina	(311)	AAQVWSQ	ECKTIV	SQYGO	ILDL	LIAET	QPKKI	CSQVGL	CTF	BD	TRGV	SAGRS	IVSDDE	PPVKS
AtAP1	(320)	AAQVWSQ	ECKTIV	SQYGO	ILDL	LIAET	QPKKI	CSQVGL	CTF	BD	TRGV	SAGRS	IVSDDE	PPVKS
SoyAP1	(317)	AAQVWSQ	ECKTIV	SQYGO	ILDL	LIAET	QPKKI	CSQVGL	CTF	BD	TRGV	SAGRS	IVSDDE	PPVKS
StAP	(298)	AAQVWSQ	ECKTIV	SQYGO	ILDL	LIAET	QPKKI	CSQVGL	CTF	BD	TRGV	SAGRS	IVSDDE	PPVKS
Cardosina A	(308)	AAQVWSQ	ECKTIV	SQYGO	ILDL	LIAET	QPKKI	CSQVGL	CTF	BD	TRGV	SAGRS	IVSDDE	PPVKS
Ciprosina B	(312)	AAQVWSQ	ECKTIV	SQYGO	ILDL	LIAET	QPKKI	CSQVGL	CTF	BD	TRGV	SAGRS	IVSDDE	PPVKS
Preprogaline A	(310)	AAQVWSQ	ECKTIV	SQYGO	ILDL	LIAET	QPKKI	CSQVGL	CTF	BD	TRGV	SAGRS	IVSDDE	PPVKS
Preprogaline B	(309)	AAQVWSQ	ECKTIV	SQYGO	ILDL	LIAET	QPKKI	CSQVGL	CTF	BD	TRGV	SAGRS	IVSDDE	PPVKS
Preprofigain A	(304)	AAQVWSQ	ECKTIV	SQYGO	ILDL	LIAET	QPKKI	CSQVGL	CTF	BD	TRGV	SAGRS	IVSDDE	PPVKS
Preprofigain B	(209)	AAQVWSQ	ECKTIV	SQYGO	ILDL	LIAET	QPKKI	CSQVGL	CTF	BD	TRGV	SAGRS	IVSDDE	PPVKS
Preprofigain C	(210)	AAQVWSQ	ECKTIV	SQYGO	ILDL	LIAET	QPKKI	CSQVGL	CTF	BD	TRGV	SAGRS	IVSDDE	PPVKS
Quimosina	(296)	AT	---	---	---	---	---	---	---	---	---	---	---	---
Fitepsina	(421)	GBSAVDC	GSILGSM	DI	EFTI	IGK	FAL	DEE	Y	ILKV	GE	AAAO	ISG	TAM
AtAP1	(430)	GBSAVDC	GSILGSM	DI	EFTI	IGK	FAL	DEE	Y	ILKV	GE	AAAO	ISG	TAM
SoyAP1	(427)	GBSAVDC	GSILGSM	DI	EFTI	IGK	FAL	DEE	Y	ILKV	GE	AAAO	ISG	TAM
StAP	(407)	GBSAVDC	GSILGSM	DI	EFTI	IGK	FAL	DEE	Y	ILKV	GE	AAAO	ISG	TAM
Cardosina A	(418)	EBQVDC	GNLSSM	EN	VSFTI	IGK	KEG	AT	Q	ISG	TAM	---	---	---
Ciprosina B	(422)	GBSAVDC	GSILGSM	DI	EFTI	IGK	FAL	DEE	Y	ILKV	GE	AAAO	ISG	TAM
Preprogaline A	(420)	GBSAVDC	GSILGSM	DI	EFTI	IGK	FAL	DEE	Y	ILKV	GE	AAAO	ISG	TAM
Preprogaline B	(419)	GBSAVDC	GSILGSM	DI	EFTI	IGK	FAL	DEE	Y	ILKV	GE	AAAO	ISG	TAM
Preprofigain A	(414)	GBSIN	GNLSSM	EN	VSFTI	IGK	KEG	AT	Q	ISG	TAM	---	---	---
Preprofigain B	(318)	GBSIN	GNLSSM	EN	VSFTI	IGK	KEG	AT	Q	ISG	TAM	---	---	---
Preprofigain C	(320)	GBSIN	GNLSSM	EN	VSFTI	IGK	KEG	AT	Q	ISG	TAM	---	---	---
Quimosina	(302)	GBSIN	GNLSSM	EN	VSFTI	IGK	KEG	AT	Q	ISG	TAM	---	---	---

Estructura primaria

La estructura primaria de las enzimas codificadas por los genes aislados en este estudio presentan los motivos que caracterizan a los zimógenos de las PAs vegetales típicas (Simões y Faro 2004): un péptido señal hidrofóbico en el extremo N-terminal, seguido de un prosegmento, y los dominios N- y C- terminal separados por un fragmento específico de plantas (PSI). En el dominio C-terminal se encuentran los residuos conservados DTG y DSG, pertenecientes al sitio activo (Figura 26). Las secuencias de aminoácidos de las tres proteasas contienen en su secuencia sitios potenciales de glicosilación. La preprogaline A presenta dos sitios de glicosilación, uno en el dominio N-terminal y otro en el C-terminal, la preprofigain A y la preprogaline B presentan un único sitio, en el dominio N-terminal y en el PSI respectivamente. La secuencia incompleta de las proteasas de *F. carica*, preprofigain B y C presentaban un potencial sitio de glicosilación en el PSI, en una localización homóloga a la presente en la preprogaline B, y la preprofigain B presentaba un sitio de glicosilación adicional en el extremo N-terminal.

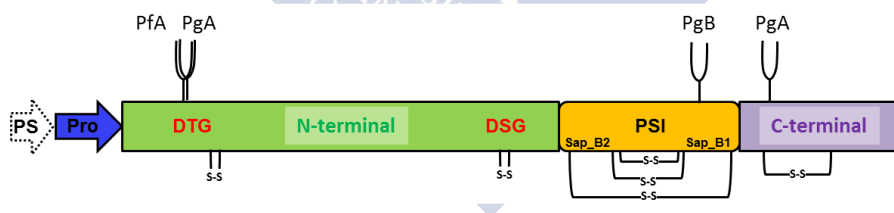


Figura 26. Representación gráfica de la estructura primaria de las PAs identificadas en este estudio. Los dominios incluyen un péptido señal (PS), un prosegmento (Pro), un dominio N-terminal (en verde) conteniendo los dos residuos catalíticos de ácido aspártico incluidos en las secuencias Asp-Thr-Gly y Asp-Ser-Gly (DTG y DSG respectivamente, en rojo), un inserto específico de plantas (PSI, en naranja), que contiene dos dominios tipo saposina B (Sap_B2 y Sap_B1) y un dominio C-terminal (en magenta). Los sitios de glicosilación están representados por una horquilla, en cuya parte superior se indica la proteasa que lo porta: preprogaline A (PgA), preprogaline B (PgB) o Preprofigain A (PfA). Los puentes disulfuro predichos están indicados mediante líneas.

El análisis de las secuencias mediante el programa CD-Search del NCBI (Marchler-Bauer et al. 2017) determinó que el PSI presentaba homología con el dominio “saposina tipo B” (Sap_B, código Prosite: PS50015) que caracteriza a la familia de proteínas denominadas SAPLIP (Bruhn 2005). Este dominio ha sido dividido en dos subdominios denominados Sap_B2 y Sap_B1 debido a la peculiaridad que presenta en las plantas, en las cuales la homología ocurre entre la secuencia N-terminal (Sap_B2; código Pfam: PF03489) y C-terminal (Sap_B1; código Pfam: PF05184) del PSI permutadas. En la Figura 27 se muestra un alineamiento del dominio “saposina tipo B” (Sap_B) de varias proteínas pertenecientes a la familia SAPLIP cuya estructura cristalográfica ha sido determinada, saposinas A-D humanas [códigos PDB: 2DOB, 1N69, 2GTG, 2RB3 (Ahn et al. 2006; Rossmann et al. 2008)], NK-Lysin (Liepinsh et al. 1997), fitepsina (código PDB: 1QDM; Kervinen et al. 1999], *S. tuberosum* (código PDB: 3RFI; Bryksa et al. 2011) y otras PAs cuya secuencia ha sido deducida a partir del ADNc, como la cardosina B y la ciprosina B (códigos Genbank: CAB40349 y CAA57510 respectivamente) procedentes del cardo y AtAP1 de *A. thaliana* (Genbank: ANM60511.1). Todas ellas presentan en este dominio 6 residuos de cisteína conservados e implicados en la formación de tres puentes disulfuro y un patrón de residuos hidrofóbicos conservados. El sitio de glicosilación, conservado en la mayoría de los PSI de las PAs típicas y en las saposinas humanas (Munford et al. 1995; Simões y Faro 2004), no está presente en la preprogaline A y preprofigain A, este sitio se encontraría entre las hélices 4 y 5 de la fitepsina (equivalente a las hélices 1 y 2 de las saposinas).

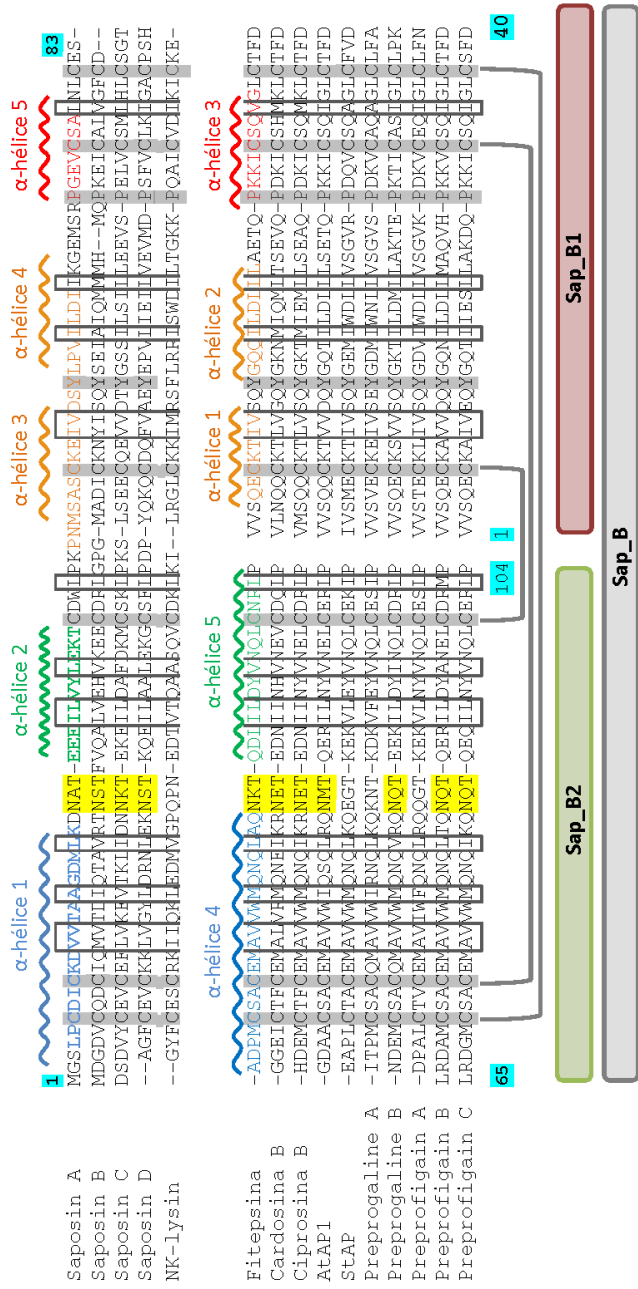


Figura 27. Alineamiento de proteínas de la familia SAPLIP (Saposinas A-D y NK-lysin) y los extremos C- y N- terminal del PSI. Las secuencias de N-glicosilación aparecen resaltadas en amarillo. Los residuos hidrofóbicos conservados aparecen en el interior de recuadros y los residuos idénticos en todas las secuencias aparecen sobre fondo gris. Los residuos coloreados pertenecientes a las hélices α son los indicados en las estructuras cristalográficas de la saposina A (Ahn et al. 2006) y de la fitepsina (Kervinen et al. 1999). La numeración (resaltada en azul) de la parte superior corresponde a la secuencia de la saposina A y la inferior al PSI de la fitepsina [nótese que la numeración del PSI se inicia en la mitad del alineamiento, los residuos que unen los dominios Sap_B1 y Sap_B2 (residuos 40-65) se han omitido]. Los puentes disulfuro se indican con líneas.

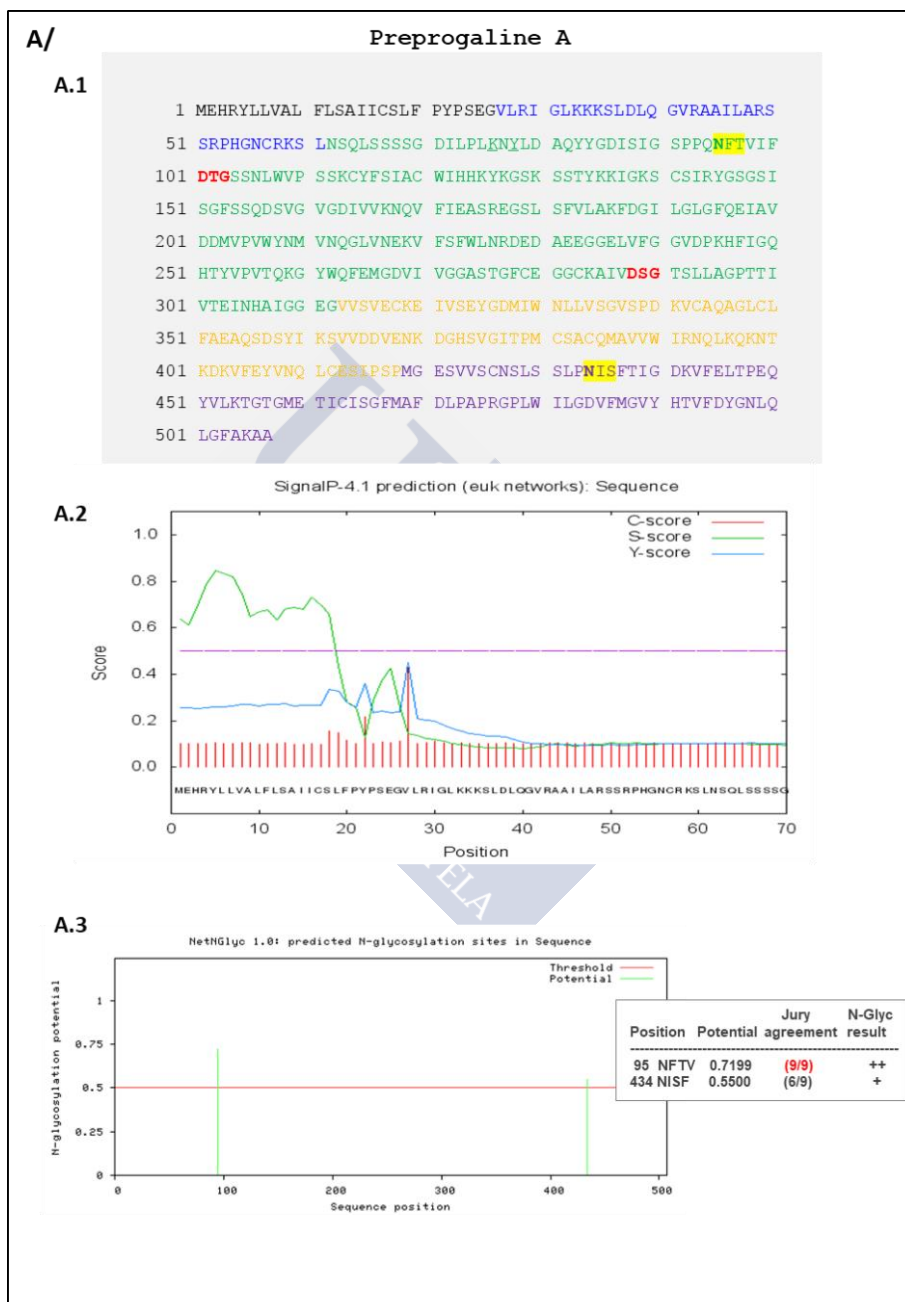
La preprogaline A, preprogaline B y preprofigain A presentan los residuos Lys/Tyr en posición equivalente a los residuos Lys11/Tyr13 presentes en la secuencia madura de la proteasa aspártica de *H. vulgare* (código Genbank CAA39602.1, corresponden a los residuos 77 y 79 de la enzima sin procesar) (indicados mediante asteriscos en la Figura 25) sugiriendo que podrían presentar un mecanismo de inactivación similar al propuesto por Kervinen y colaboradores (1999) para la fitepsina, en el cual el sitio activo está bloqueado por el prosegmento y por los 13 residuos del extremo N-terminal de la enzima madura, el anclaje tendría lugar mediante interacciones iónicas establecidas entre los residuos Lys/Tyr y los residuos catalíticos de ácido aspártico (Kervinen et al. 1999) (Vairo-Cavalli et al. 2013). Los residuos implicados en la inactivación se encontrarían en las posiciones Lys76/Tyr78, Lys75/Tyr77 y Lys70/Tyr72 de las secuencias de la preprogaline A, preprogaline B y preprofigain A respectivamente.

Las tres proteasas presentan los cuatro aminoácidos conservados FAEA, en el extremo C- terminal.

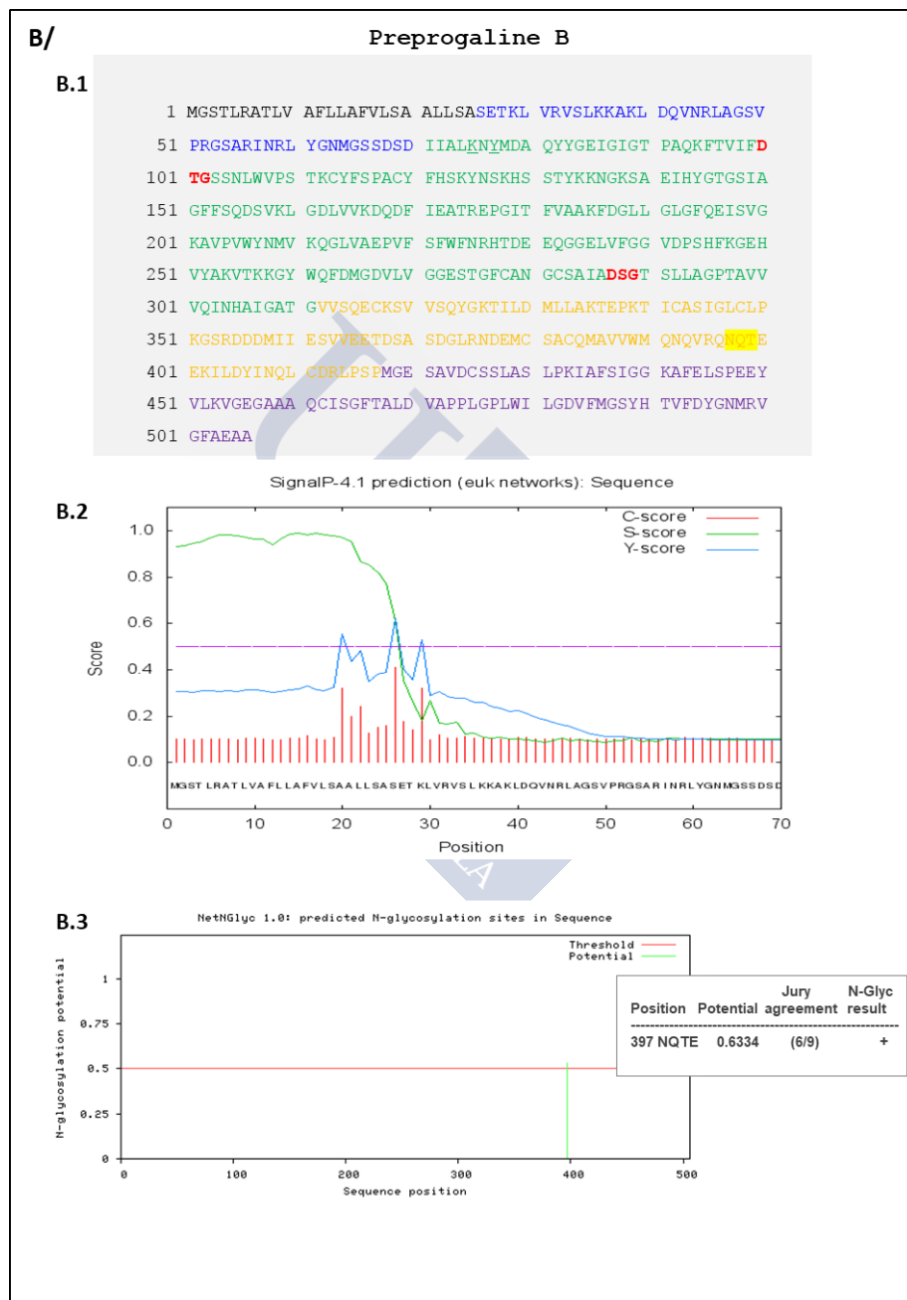
La preprogaline B y la secuencia parcial de la preprofigain C presentan el motivo KGE en posición equivalente al motivo RGD característico de la cardosina A (Simões et al. 2005).

En la Figura 28 se pueden consultar los aminoácidos de las PAs aisladas que forman parte de cada uno de los motivos anteriormente citados, así como los gráficos de predicción de los péptidos señal y de los sitios de N-glicosilación.

(Figura 28)



(Figura 28)



(Figura 28)

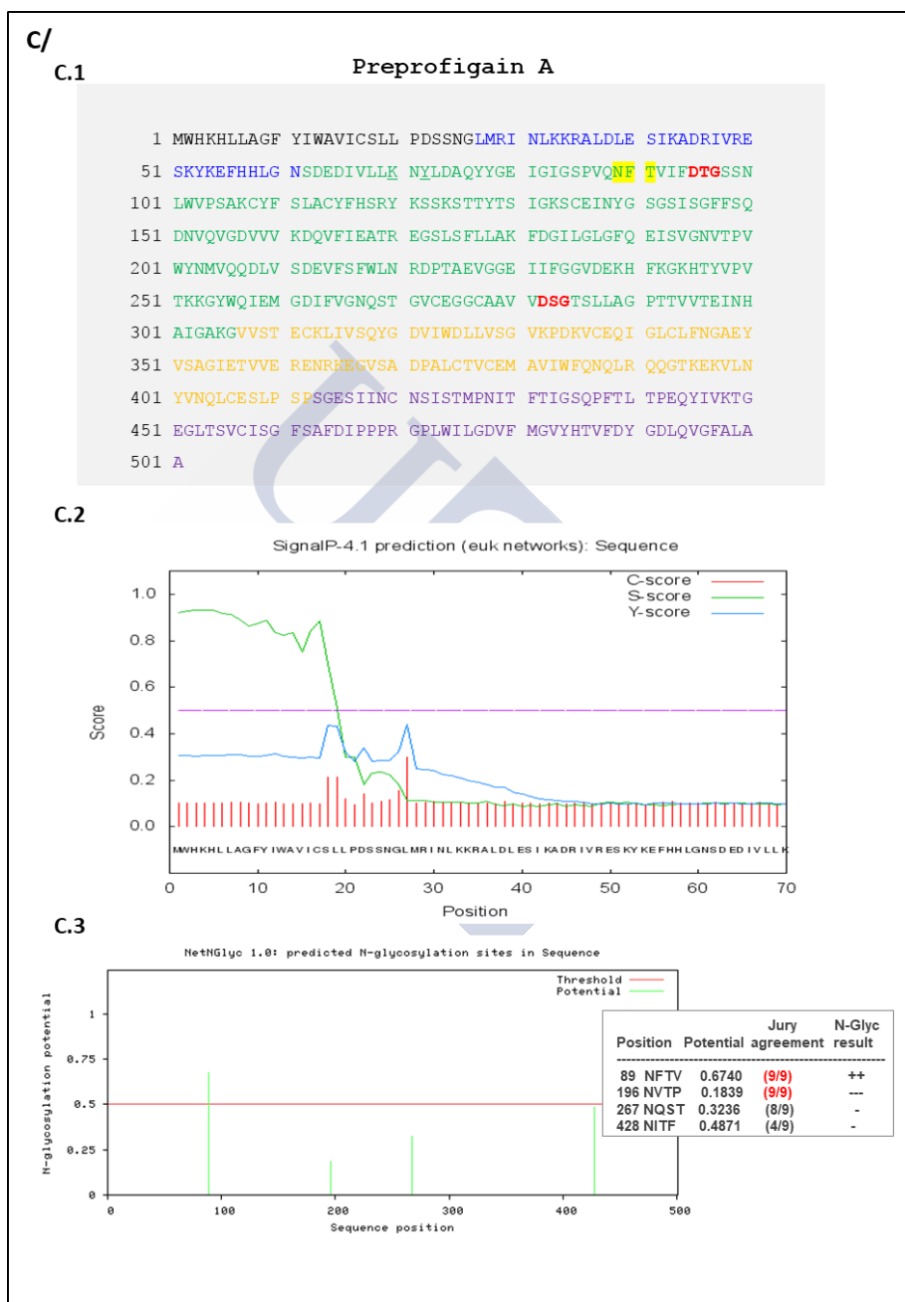


Figura 28. Análisis de las secuencias de aminoácidos de la preprogaline A (A), preprogaline B (B) y preprofigain A (C). En cada uno de los recuadros, en la parte superior, **A.1, B.1 y C.1**, se muestra la secuencia de aminoácidos deducida a partir del ADNc para las secuencias *gv1AP*, *gv2AP* y *fc1AP* respectivamente. En negro aparecen los residuos que pertenecen al péptido señal, en azul los residuos del prosegmento, en verde los residuos del dominio N-terminal y en magenta los pertenecientes al dominio C-terminal, interrumpidos por los residuos pertenecientes al PSI, que aparecen en naranja. Los posibles sitios de glicosilación se muestran sobre fondo amarillo, los residuos que forman parte del sitio activo se muestran en color rojo, y subrayados los residuos Lys/Tyr posiblemente implicados en el mecanismo de inactivación de las proteasas. En la parte intermedia, **A.2, B.2 y C.2**, gráficos de predicción de los péptidos señal realizados con SignalP 4.1 Server. Las líneas rojas (C-score) indican la probabilidad de que un péptido señal se procese en ese punto, la línea verde (S-score), indica la probabilidad de que un aminoácido pertenezca al péptido señal, y la línea azul (Y-score) es una combinación de las anteriores, para que sea máxima en aquellos aminoácidos que tengan una alta probabilidad de ser un sitio de corte. En la parte inferior, **A.3, B.3 y C.3**, gráficos de predicción de sitios de N-glicosilación realizados con NetNGlyc 1.0 Server. El eje X representa la longitud de la proteína desde el extremo N- al C-terminal, un residuo con un “potencial” (línea verde) que atravesase la línea horizontal roja se predice que está N-glicosilado. En el recuadro se muestran los datos de predicción del programa.

Estructura tridimensional

Hoy en día se conoce únicamente la estructura cristalina de dos PAs vegetales, la forma inmadura sin el péptido señal de la fitepsina (Kervinen et al. 1999) y la forma procesada de la cardosina A procedente del cardo (Frazão et al. 1999). Estas estructuras nos han permitido disponer de un modelo para realizar una aproximación teórica de la estructura tridimensional de las PAs identificadas en este estudio. El programa CPHmodels 3.0 seleccionó, como era de esperar, en base a su homología la fitepsina (código PDB: 1QDM) para realizar el análisis. Las estructuras proteicas modelizadas fueron analizadas mediante el programa Pymol (Delano Scientific LLC 2006). Las tres proteasas presentan una estructura secundaria similar, en la Figura 29 se muestra la estructura de la preprogaline B. La forma madura consta fundamentalmente de láminas plegadas β y presenta

dos dominios homólogos, con el sitio activo localizado en una hendidura situada entre ambos, cada dominio contribuiría con uno de los motivos de la secuencia catalítica Asp-Thr-Gly (DTG) y Asp-Ser-Gly (DSG). El PSI constituye una subunidad independiente en la estructura, está situado en el dominio C-terminal y constituido por cinco hélices α .

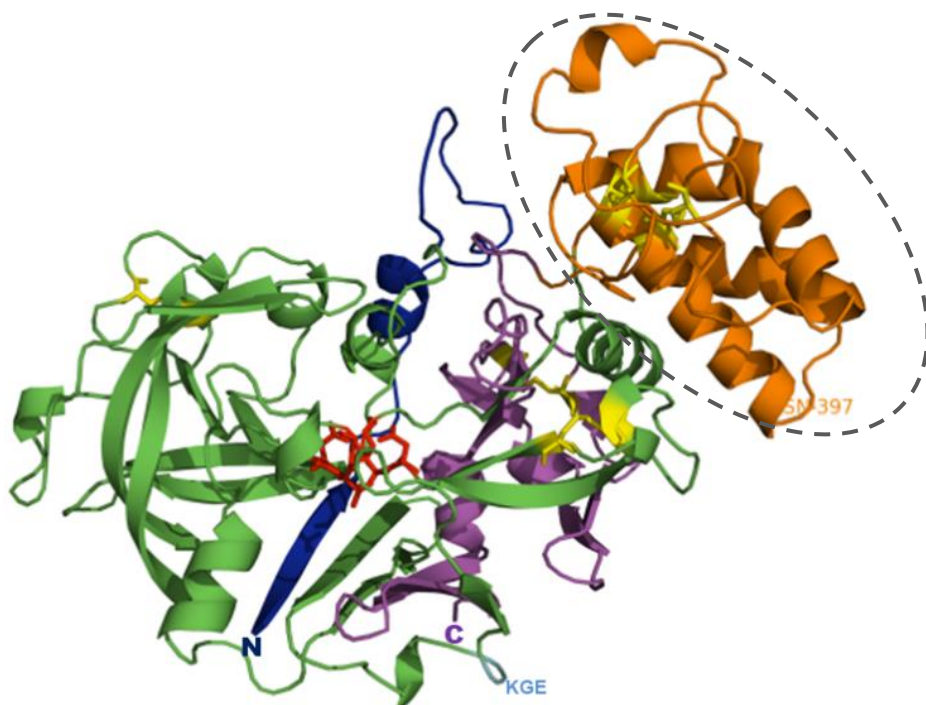


Figura 29. Modelo tridimensional de la preprogalina B. El prosegmento se muestra en azul, la cadena pesada (HC) se muestra en verde, el PSI en naranja (resaltado con un círculo discontinuo) y la cadena ligera (LC) en color magenta. Los residuos de las tríadas catalíticas DTG y DSG aparecen como barras rojas, los residuos de cisteína que participan en los enlaces disulfuro se muestran como barras amarillas, el residuo de asparagina posiblemente implicado en la N-glicosilación se muestra como ASN-397 y en azul celeste el motivo KGE. La representación fue creada utilizando el software Pymol (DeLano Scientific LLC).

Como otros miembros de la familia de las PAs vegetales típicas contiene 6 residuos de cisteína implicados en la formación de tres puentes disulfuro. Otros tres puentes disulfuro se encontrarían fuera del PSI (Tabla 14).

Tabla 14. Predicción de los residuos de cisteína implicados en la formación de puentes disulfuro.

Proteasa	Residuos Cys Implicados	Localización
Preprogaline A	114-120	Cadena pesada
	279-283	Cadena pesada
	318-412	PSI
	343-384	PSI
	349-381	PSI
	426-463	Cadena ligera
Preprogaline B	113-119	Cadena pesada
	278-282	Cadena pesada
	317-411	PSI
	342-383	PSI
	348-380	PSI
	425-462	Cadena ligera
Preprofigain A	108-114	Cadena pesada
	273-277	Cadena pesada
	312-406	PSI
	337-378	PSI
	343-375	PSI
	429-457	Cadena ligera

La validación del modelo tridimensional se realizó utilizando el programa ProSA (Wiederstein y Sippl 2007). Este programa reconoce errores en la estructura tridimensional mediante el cálculo del parámetro Z-score, que indica la calidad del modelo general, y mediante la desviación de la energía total de la estructura obtenida con respecto a la distribución de la energía derivada de conformaciones al azar (Figura 30).

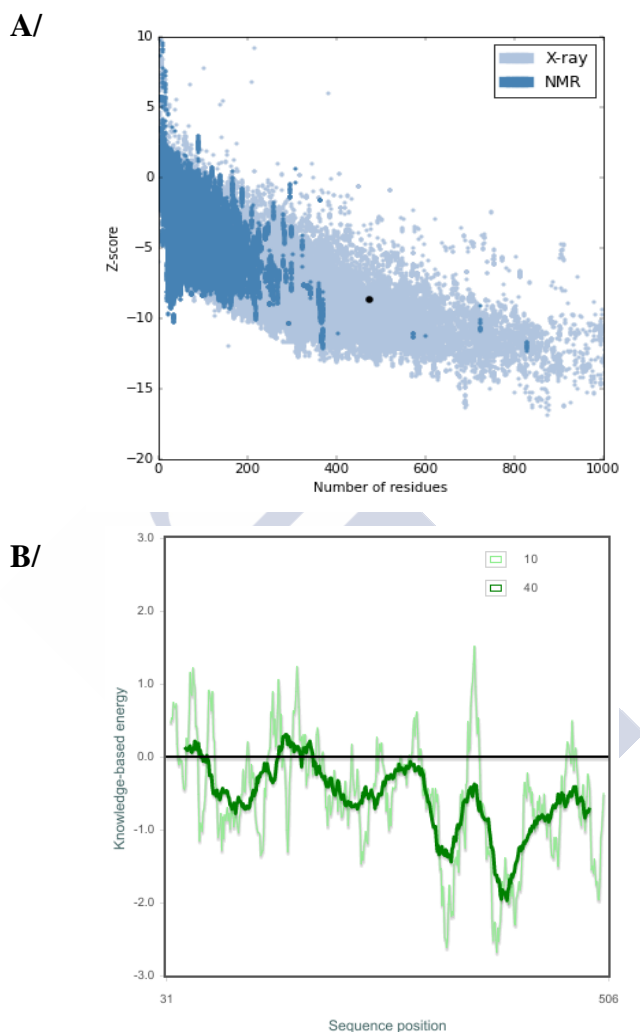


Figura 30. Resultados de validación del modelo 3D de la preprogaline B mediante el programa ProSA. (A) Gráfico en el que se muestra el Z-score de las proteínas de la base de datos PDB cuyas estructuras han sido determinadas experimentalmente por cristalografía de rayos X (azul claro) o por resonancia magnética nuclear (azul oscuro) con respecto a su longitud. El gráfico muestra únicamente aquellas con cadenas de aminoácidos inferiores a 1000 residuos y un Z-score inferior a 10. El Z-score del modelo generado para la preprogaline B se indica con un punto negro. **(B)** Gráfico de energía de los aminoácidos del modelo 3D de la preprogaline B.

En la Figura 30.A puede observarse el gráfico en el que se relacionan las puntuaciones Z-score de otras proteínas cuya estructura ha sido determinada experimentalmente y el modelo generado. El Z-score obtenido para el modelo de la preprogaline B fue de -8.69 (indicado con un punto negro en la figura), este valor se encuentra dentro del rango de todas las proteínas nativas resueltas experimentalmente, y específicamente se encuentra en el rango de las resueltas por cristalografía de rayos X.

El gráfico de la Figura 30.B muestra la calidad del modelo representando el perfil de energía en función de la posición de los aminoácidos de la preprogaline B. En general, los valores que sobrepasan la línea horizontal (valores positivos) corresponden a partes problemáticas o erróneas de la estructura que se está analizando. El gráfico con ventanas de 10 residuos (color verde claro) contiene grandes fluctuaciones y es de valor limitado para evaluar el modelo, por ello el gráfico es “atenuado” calculando la energía media sobre fragmentos de 40 residuos de aminoácidos (color verde oscuro). En el modelo de la preprogaline B se observan muy pocos residuos de aminoácidos problemáticos o erróneos en el modelo tridimensional obtenido, ya que los valores fluctúan bajo el cero. Por lo tanto, se vuelven a validar como estables los residuos de aminoácidos del modelo.

Los modelos generados para la preprogaline A y preprofigain A también fueron validados, presentando un Z-score de -8.36 y -7.73 respectivamente.

4.2.2.2 Análisis filogenético

Las secuencias de aminoácidos de la preprogaline A, preprogaline B y preprofigain A fueron comparadas con las de otras PAs vegetales mediante un análisis filogenético realizado con el programa MEGA 6.0, utilizando el método de máxima verosimilitud (MV) (Figura 31).

Todas las secuencias utilizadas en la elaboración del árbol corresponden a PAs de plantas sin procesar, clasificadas como típicas en función de su estructura, y de las cuales se tienen evidencias experimentales, no únicamente una predicción bioinformática. Como grupo externo para enraizar el árbol se utilizó la chlpsin, una PA del alga verde unicelular *Chlamydomonas reinhardtii* (Almeida et al. 2012). El modelo de sustitución de aminoácidos que mejor se ajustó para realizar el análisis fue el WAG (Whelan And Goldman model) con una distribución gamma y sitios invariantes (Whelan y Goldman 2001).

En base al árbol filogenético obtenido se podrían definir dos grandes grupos dentro de las PAs analizadas (Figura 31, I y II).

El grupo I incluiría cuatro subgrupos (Ia₁, Ia₂, Ib₁ y Ib₂). El grupo Ia₁ comprende PAs de leguminosas como el caupí (*Vigna unguiculata*), la judía común (*Phaseolus vulgaris*), la PA soyAP1 aislada desde la soja (*Glycine max*) y otras PAs procedentes de no leguminosas, como dos especies de la familia de las brasicáceas, el nabo (*Brassica napus*) y la procedente del organismo modelo *A. thaliana*. No han sido atribuidas funciones específicas para estas enzimas, pero se ha descrito que la PA del caupí podría estar implicada en procesos de senescencia en hojas (Cruz de Carvalho et al. 2004) y la soyAP1, aislada de la soja estaría implicada en la muerte celular programada durante la autólisis de elementos traqueales inmaduros y células del tubo criboso, participando en la maduración del sistema vascular (Terauchi et al. 2004). El grupo Ia₂ englobaría a la mayoría de las PAs procedentes de especies monocotiledóneas utilizadas para elaborar el árbol. Así, este grupo incluye una PA de la piña (*Ananas comosus*), la aislada del arroz (*O. sativa*), la fitepsina (*H. vulgare*) y la WAP1 aislada del trigo (*Triticum aestivum*), todas ellas pertenecientes al orden Poales. En cuanto a la función de las proteasas pertenecientes a este grupo, la oryzasin se expresa durante la maduración y germinación de las semillas, por lo que se cree está implicada en eventos que ocurren

durante este periodo (Asakura et al. 1995) y la fitepsina, al igual que la soyAP1 perteneciente al grupo Ia₁, estaría implicada en el desarrollo del sistema vascular (Runeberg-Roos y Saarma 1998). El grupo Ib incluiría al subgrupo Ib₁, formado únicamente por la preprogaline B aislada en este estudio y el subgrupo Ib₂, constituido por PAs procedentes de especies pertenecientes a la familia Asteraceae, como son: la cardosina A (Faro et al. 1999), cardosina B (Vieira et al. 2001), ciprosina B (Cordeiro et al. 1998), cenprosin (Domingos et al. 2000) y preprosilpepsin 1. Estas enzimas han sido aisladas sobre todo desde flores y se ha propuesto que participan en eventos relacionados con la floración, son las más estudiadas dentro de las que figuran en el árbol debido a su capacidad para coagular la leche (Shah et al. 2014). Finalmente, el grupo II, que incluiría el subgrupo IIa, al que pertenecería la WAP2 aislada del trigo (única monocotiledónea que figura fuera del grupo Ia₂) y el subgrupo IIb compuesto por la PAs, preprogaline A y preprofigain A aisladas en este estudio, la soyAP2 (Terauchi et al. 2004), una PA aislada del tabaco y tres PAs aisladas desde las hojas de la patata dulce (*Ipomoea batatas*) (Chen et al. 2015) y de dos especies de la familia de las *Solanaceas*, el tomate y la patata. Estas dos últimas enzimas parece que se producen en respuesta a estrés biótico (Guevara et al. 2005; Schaller y Ryan 1996). La secuencia de la chlapsin, utilizada como grupo externo, aparece en una rama aislada. El análisis filogenético no nos permite establecer una hipótesis sobre la posible función de las PAs aisladas en este estudio.

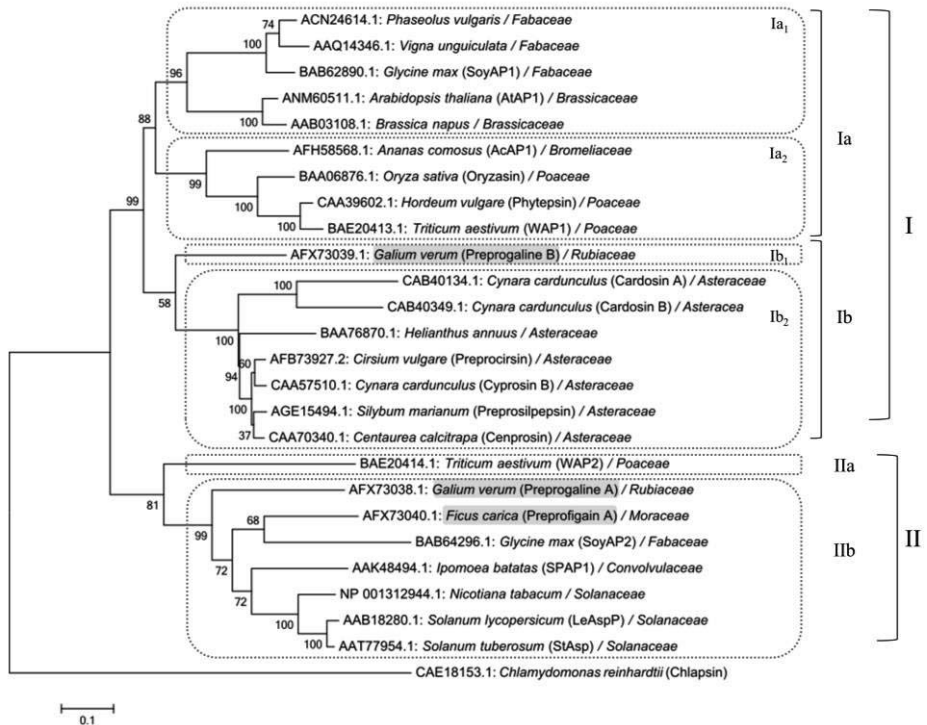


Figura 31. Árbol filogenético de la preprogaline A, preprogaline B y preprofigain A y otras PAs vegetales típicas. Los números de acceso de las secuencias se indican en cada rama, seguidos de la especie y familia taxonómica a la que pertenecen; entre paréntesis se indica el nombre vulgar de alguna de las proteasas. Las PAs objeto de esta tesis aparecen sombreadas en gris. Los grupos establecidos en base al árbol se indican en el margen derecho de la imagen. El análisis filogenético se realizó utilizando el modelo de máxima verosimilitud con el programa MEGA 6 (Tamura et al 2013). La fiabilidad del árbol filogenético inferido se realizó mediante un análisis *bootstrap* con 500 réplicas, estos valores fueron convertidos en porcentajes y se muestran en los nodos. Las longitudes de las ramas horizontales se encuentran a escala y son proporcionales al porcentaje de divergencia. La barra de escala representa 0.1 variaciones de aminoácidos por sitio. Como grupo externo se utilizó una PA del alga unicelular *Chlamydomonas reinhardtii*.

4.3 Expresión de las proteasas aspárticas en *E. coli*

El sistema de expresión seleccionado inicialmente para la producción de las PAs fue el constituido por vectores de la serie pET y como hospedadores las cepas de *E. coli* BL21(DE3) o Rosetta-gami (DE3).

4.3.1 Expresión de la progaline A en el vector pET30a(+)

La secuencia codificante de la preprogaline A sin la parte correspondiente al péptido señal, que denominaremos progaline A, fue clonada en el vector de expresión pET30a(+). En este vector se realizó una fusión traduccional que constaba de una cola de histidinas y un sitio de corte para la enterocinasa en el extremo N-terminal, fusionada con la progaline A (Figura 32). El peso molecular estimado de la proteasa recombinante sería de 57.9 kDa.

```

1  MHHHHHSSG LVPRGSGMKE TAAKFERQH MDSPDLGTDD DDKAMADIGS
51  EFVLRIGLKK KSLDLQGVRA AILARSSRPB GNCRKSLNSQ LSSSSGDILP
101 LKNYLDAQYY GDISIGSPPQ NFTVIFDTGS SNLWVPSSKC YFSIACWIHH
151 KYKGSKSSTY KKIGKSCSIR YGSGSISGFS SQDSVGVGDI VVKNQVFIEA
201 SREGSLSFVL AKFDGILGLG FQEIAVDDMV PVWYNMVNQG LVNEKVFSEW
251 LNRDEDAEEG GELVFGGVDP KHFIGQHTYV PVTQKGYWQF EMGDVIVGGA
301 STGFCEGGCK AIVDSGTSLL AGPTTIVTEI NHAIGGEGVV SVECKEIVSE
351 YGDMIWNLLV SGVSPDKVCA QAGLCLFAEA QSDSYIKSVV DDVENKDGHS
401 VGITPMCSAC QMAVVWIRNQ LKQKNTKDKV FEYVNQLCES IPSPMGESVV
451 SCNSLSSLPN ISFTIGDKIF ELTPEQYVLK TGTGMETICI SGFMAFDLPA
501 PRGPLWILGD VFMGVYHTVF DYGNLQIGFA KAA*

```

Figura 32. Secuencia de aminoácidos de la proteína producto de la expresión de la construcción pET30a-progv1AP. Los aminoácidos subrayados corresponden a los aportados por el vector, en rojo se muestra la cola de histidinas y en verde un sitio de corte para la enzima enterocinasa. Los residuos pertenecientes al prosegmento de la progaline A se muestran en azul.

La inducción de la cepa de expresión BL21(DE3) se realizó con una concentración IPTG 1 mM durante 4 y 20 horas, transcurridas las cuales se realizó la extracción del contenido proteico celular total y de la fracción citoplasmática soluble e insoluble (Figura 33).

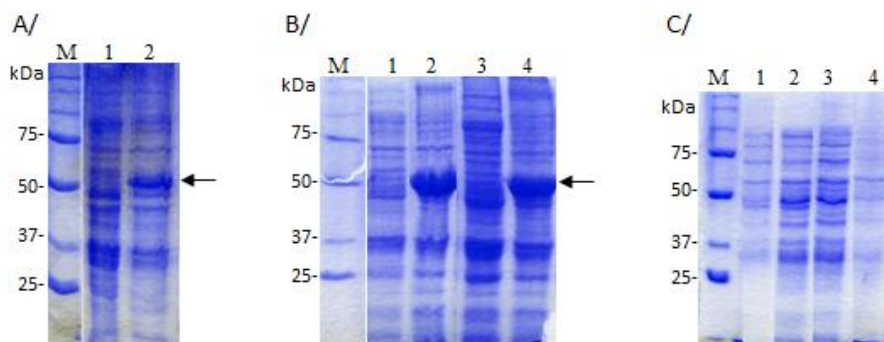


Figura 33. Análisis mediante SDS-PAGE de la expresión de la progaline A en *E. coli* utilizando el vector pET30a(+). (A) Fracción proteica celular total: cultivo inducido con IPTG durante 4 horas (calle2) y sin inducción (calle 1). Fracción proteica citoplasmática insoluble (B) y soluble (C): en ambos geles se la calle 2 corresponde a un cultivo inducido con IPTG durante 4 horas y la calle 4 durante 20 horas, las calles 1 y 3 corresponden a los controles negativos respectivos.

El sistema de expresión permitió la obtención de la progaline A (Figura 33.A). Así, cuando se realizó la extracción de proteínas de la cepa recombinante se observaba una banda de un peso molecular aproximado de 58 kDa (calle 2), que se correspondía con el tamaño estimado de la proteasa recombinante no glicosilada. Esta banda no estaba presente en las células cultivadas en ausencia del inductor de expresión (calle 1).

La progaline A se expresaba en forma de agregados citoplasmáticos insolubles, también llamados cuerpos de inclusión (Figura 33.B), que no estaban presentes en la fase citoplasmática soluble (Figura 33.C). En la figura 33.B puede observarse la banda correspondiente a la progaline A en los cultivos inducidos con IPTG durante 4 o 20 horas respectivamente (calles 2 y 4), la cual no aparece en las células no inducidas (calles 1 y 3). Cuando se analizó la fracción soluble del citoplasma, el patrón de bandas de las células

incubadas en presencia o ausencia del inductor IPTG era el mismo (Figura 33.C), no apreciándose ninguna banda adicional que pudiese corresponder a la proteasa recombinante. Se modificó el protocolo de obtención de la fracción soluble añadiendo el detergente no desnaturalizante CHAPS al tampón de lisis, para solubilizar la proteasa en caso de que estuviese asociada a la membrana, la banda correspondiente a la progaline A seguía sin apreciarse cuando la fracción soluble fue analizada mediante SDS-PAGE.

La producción de proteínas heterólogas en *E. coli* conduce en muchas ocasiones a la obtención de la proteína recombinante mal plegada, en forma de cuerpos de inclusión, sin embargo, una proporción de la misma está normalmente soluble en el interior celular. Se han desarrollado varios métodos para aumentar la proporción de la proteína recombinante soluble en relación con la que se encuentra formando parte de agregados. En general, condiciones que disminuyan la tasa de síntesis, como son bajas temperaturas de inducción o el crecimiento en medios mínimos, tienden a aumentar el porcentaje de la proteína expresada que se encuentra en forma soluble (Sørensen y Mortensen 2005). Para intentar producir la progaline A en forma soluble se modificó la temperatura y tiempo de inducción, que pasó de ser 37 °C durante 4 horas (Figura 34.A) a 20 °C durante 18 horas (Figura 34.B). El cultivo a baja temperatura no aumentó la proporción de la proteína soluble en el citoplasma celular, o por lo menos no lo hizo de forma apreciable, por ello abordamos la purificación y “replegamiento” de la progaline A a partir de los cuerpos de inclusión.

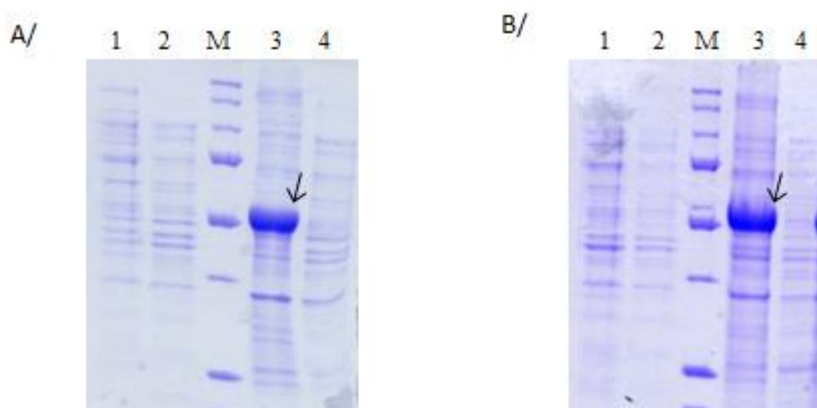


Figura 34. Influencia de la temperatura de inducción en la producción de la progaline A en *E. coli* BL21(DE3). Proteínas obtenidas a partir de un cultivo incubado 4 horas a 37 °C (A) o 18 horas a 20 °C (B). Las calles 1 y 2 corresponden a muestras del citoplasma soluble y las calles 3 y 4 del citoplasma insoluble de un cultivo inducido IPTG 1mM (calles 1 y 3) y un cultivo control no inducido (calles 2 y 4). M: marcador de pesos moleculares. La banda correspondiente a la progaline A se indica con una flecha.

4.3.1.1 Purificación de los cuerpos de inclusión

Los cuerpos de inclusión fueron purificados desde el lisado celular en condiciones desnaturalizantes mediante cromatografía de afinidad con iones de níquel inmovilizados (IMAC). Los extractos obtenidos fueron desnaturalizados en presencia de urea y la cola de poli-histidinas de la secuencia N-terminal de la progaline A recombinante permitió la purificación de esta debido a su afinidad por los iones níquel. Se abordó la optimización de la concentración de imidazol, sustancia análoga a la histidina que permite la elución de la proteína recombinante. La pureza y cantidad de proteína recombinante obtenida era prácticamente la misma con las tres concentraciones de imidazol testadas para la elución (0.25 M, 0.5 M e 0.75 M) (Figura 35).

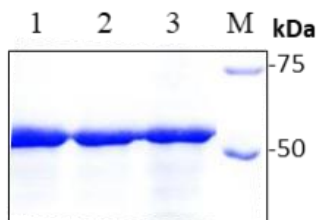


Figura 35. SDS-PAGE para establecer la concentración de imidazol empleada para la purificación mediante IMAC de cuerpos de inclusión producidos en *E. coli* BL21(DE3) transformada con la construcción pET30a-progv1AP. La elución se realizó utilizando diferentes concentraciones de imidazol: 0.25 M (calle1), 0.50 M (calle2), 0.75 M (calle3). M: marcador pesos moleculares, indicados en margen derecho de la imagen.

4.3.1.2 Plegado y activación de la progaline A

La desnaturalización mediante urea de los cuerpos de inclusión para su purificación conduce a la pérdida de la actividad enzimática, para poder renaturalizar la progaline A a un estado similar al nativo es necesario la eliminación del agente desnaturalizante (Singh y Panda 2005). Se abordaron dos protocolos diferentes basados en estudios previos realizados con (i) la proteasa aspártica cardosina A de *C. cardunculus* (Castanheira et al. 2005) y (ii) un método de plegado universal patentado (Lin 2006).

(i) Los cuerpos de inclusión solubilizados en urea, conteniendo el agente reductor β -ME, fueron renaturalizados mediante dilución rápida en Tris base y ajuste del pH de la solución con HCl a pH 8. Cuando la renaturalización se realizó a temperatura ambiente la proteasa era procesada, observándose una banda de 34 kDa en el análisis mediante SDS-PAGE, que por su peso molecular podría corresponder a la cadena pesada de la progaline A sin el prosegmento y sin la secuencia de histidinas (Figura 36, calle 6); el proceso de renaturalización a 4 °C no daba lugar a cambios en el tamaño molecular de la progaline A (calle 3).

La incubación de las muestras renaturalizadas mediante ambos métodos a pH ácido, para su activación, no dio lugar a cambios en el patrón de bandas observado mediante SDS-PAGE (calles 1, 2, 4 y 5).

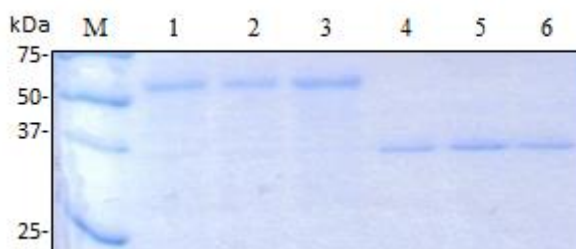


Figura 36. SDS-PAGE de cuerpos de inclusión de la progaline A plegados mediante el método de dilución rápida. Los cuerpos de inclusión fueron plegados a 4°C (calle 1-3) y a temperatura ambiente (calles 4-6). Después del proceso de plegado la progaline B fue incubada a pH 3 (calles 1 y 4) y a pH 4 (calles 2 y 5) para su activación.

Las muestras replegadas a 4 °C o a temperatura ambiente e incubadas a diferentes pHs no mostraron actividad cuando se utilizó como sustrato κ -caseína o BSA. En la Figura 37 puede apreciarse como la BSA (~66 kDa) (calle 4) permanecía intacta cuando se incubó con las muestras de progaline A replegadas a temperatura ambiente (calles 1-3), únicamente se observaban una banda adicional que correspondía al fragmento de 34 kDa de la progaline A.

(ii) Se modificó el protocolo de renaturalización de los cuerpos de inclusión basándose en otro método de dilución, método de replegado universal (Lin 2006). A diferencia del método anterior el tampón en el cual se resuspenden los cuerpos de inclusión contiene, además de urea y el agente reductor β -ME, EDTA, glicina, DTT y glutatión oxidado y reducido. Una vez replegadas las muestras estas se incubaron en tampones a diferentes pHs para su activación. Las muestras sometidas a este proceso de renaturalización tampoco

presentaron actividad enzimática cuando se utilizó κ -caseína o BSA como sustrato.

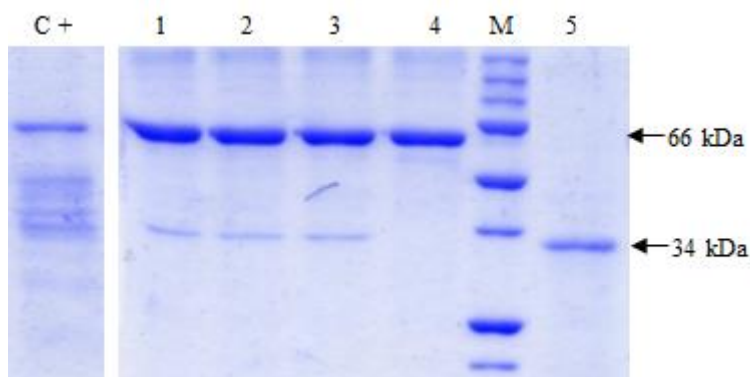


Figura 37. Análisis de la actividad proteolítica de la progaline A replegada mediante el método de dilución rápida a temperatura ambiente. La digestión se realizó a 37 °C durante 2 horas utilizando BSA como sustrato. En las calles 1 y 2 la proteína replegada fue preincubada con tampón a pH 3 y 4 a 37 °C toda la noche antes de realizar la digestión; en la calle 3 la proteasa replegada se añadió sin preincubación previa. Las calles 4 y 5 corresponden al sustrato BSA (66 kDa) y a la progaline A (34 kDa) replegada respectivamente. En la calle C+, control positivo utilizando Chymax® Plus.

4.3.2 Expresión de la progaline A y la profigain A en el vector pET21a(+)

La expresión de la preprogaline A en el vector pET30a(+), utilizando la cepa BL21(DE3), daba lugar a la producción de la proteína inactiva en forma de cuerpos de inclusión, por lo que se abordó una nueva estrategia para obtener las proteasas en su forma activa. Una de las limitaciones para producir proteínas correctamente plegadas en *E. coli* es el alto potencial reductor en el compartimento citoplasmático, ya que los puentes disulfuro se forman normalmente en el espacio periplásmico (Ke y Berkmen 2014). Teniendo esto en cuenta se seleccionó la cepa de *E. coli* Rosetta-gami (DE3) que presenta dos mutaciones, gor (glutación reductasa) y trxB (tiorredoxin reductasa), que favorecen la formación de puentes disulfuro en su

citoplasma. Además, provee ARNt para siete codones que podrían evitar problemas de transcripción derivados de las diferencias en el pool de los ARNt entre *E. coli* y *F. carica* o *G. verum*. Esta cepa presenta resistencia a varios antibióticos, entre ellos la kanamicina, de ahí que no fuese posible utilizar la construcción pET30a-progv1AP previamente utilizada para la transformación de BL21(DE3). El vector seleccionado para la expresión fue pET21a(+).

La secuencia codificante de la preprogaline A y de la preprofigain A sin el péptido señal se insertó en el sitio de restricción NheI, de forma que la secuencia de las proteasas recombinantes presentaba los aminoácidos Met-Ala-Ser en el extremo N-terminal, codificados por el vector, seguidos de la secuencia de las proteasas sin el prosegmento; en este caso las proteasas recombinantes no portaban en su secuencia la “cola” de histidinas para su purificación mediante iones níquel.

La inducción mediante IPTG 1 mM durante 4 horas de las cepas portando la construcción pET21a-progv1AP o pET21a-profc1AP daba lugar en ambos casos a la producción de una proteína de aproximadamente 52 kDa que era recuperada en la fracción citoplasmática insoluble. El tamaño de las proteínas producidas se corresponde con el peso molecular teórico del producto de expresión de pET21a-progv1AP (progaline A, 52.4 kDa) y de pET21a-profc1AP (profigain A, 52.2 kDa).

Para intentar producir las proteasas recombinantes en forma soluble se ensayaron diferentes temperaturas de cultivo (20 °C, 30 °C y 37 °C) durante 14 horas de inducción con IPTG 1 mM. El análisis mediante SDS-PAGE no evidenció una mayor producción de proteasa recombinante en forma soluble en ninguna de las temperaturas ensayadas.

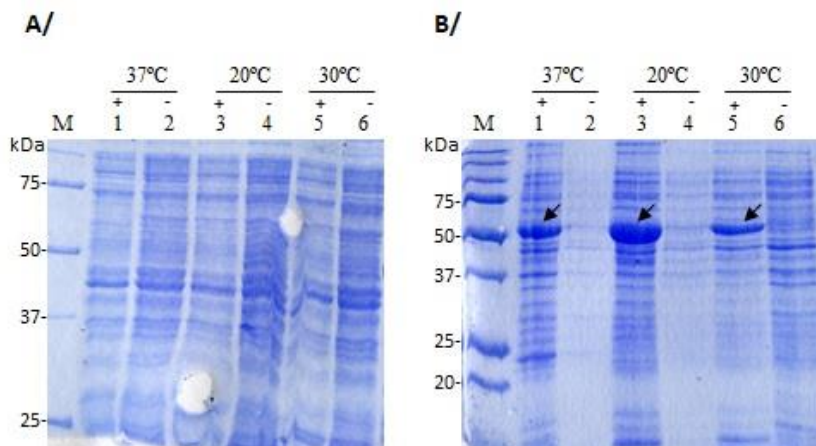


Figura 38. Análisis de la expresión de la profigain A a diferentes temperaturas de cultivo. SDS-PAGE de la fracción citoplasmática soluble (A) e insoluble (B) de la cepa Rosetta-gami (DE3) transformada con la construcción pET21a-profc1AP inducida con IPTG 1mM (calles 1,3 y 5, indicadas con símbolo +), o sin presencia del inductor (calles 2, 4 y 6, indicadas con símbolo -) durante 14 horas de cultivo a 37 °C (calles 1 y 2), 20 °C (calles 3 y 4), 30 °C (calles 5 y 6). M: marcador de pesos moleculares. Las flechas en el gel B/ indican la banda que correspondería a la profigain A.

En la Figura 38 puede apreciarse como la profigain A se producía mayoritariamente como cuerpos de inclusión en los cultivos inducidos con IPTG (gel B/, calles 1, 3 y 5); en la fracción citoplasmática soluble no se observaba ninguna banda que pudiese corresponder a la profigain A en los cultivos inducidos a las diferentes temperaturas ensayadas (gel A/ calles 1, 3 y 5).

A pesar de ello se analizó la actividad κ -caseinolítica del citoplasma soluble de ambas cepas incubadas a 20 °C en presencia y ausencia del inductor. Para ello el citoplasma soluble fue dializado frente a tampón Tris HCl 20 mM pH 8, y la concentración de proteínas fue ajustada a 10 mg/ml. Las muestras se sometieron a un proceso de activación mediante incubación con tampones a diferentes pHs (2.5; 3; 4 y 5) durante 2 horas a 37 °C , transcurridas la cuales se

analizó su actividad κ -caseinolítica. Se comprobó que no existían diferencias significativas entre las muestras incubadas en presencia o ausencia del inductor.

En vista de los resultados se abordó el plegado de los cuerpos de inclusión mediante el método de plegado universal (Lin 2006) utilizado anteriormente para replegar la progaline A producida en BL21(DE3), pero en este caso los cuerpos de inclusión no fueron purificados, sino que se utilizó la fracción citoplasmática insoluble total. Después del proceso de plegado y diálisis las muestras se incubaron a diferentes pHs (2.5; 3; 4 y 5) durante 2 horas a 37 °C con la finalidad de activar las proteasas recombinantes. No se apreciaron diferencias significativas en la actividad proteolítica sobre la κ -caseína entre las muestras incubadas en presencia o ausencia del inductor.

4.4 Expresión de las proteasas aspárticas en *Pichia pastoris*

Debido a la imposibilidad de obtener las proteasas recombinantes en su forma activa empleando un hospedador procariótico abordamos su expresión en el hospedador eucariota *P. pastoris*.

Las proteasas aspárticas preprogaline A, preprogaline B y preprofigain A fueron clonadas bajo un promotor inducible, utilizando el vector pPICZ α o/y bajo un promotor constitutivo, utilizando el vector pGAPZ α . Todos los genes se clonaron sustituyendo el fragmento correspondiente al péptido señal nativo por el factor- α de *S. cerevisiae* (9.3 kDa) aportado por el vector, que facilita la secreción de la proteína recombinante al exterior celular. Las proteasas expresadas se denominaron omitiendo el prefijo “pre” que hace referencia al péptido señal nativo. En el caso del gen preprofigain A se realizó una construcción en la que se eliminó también el prosegmento, la proteína así expresada la denominamos figain A.

En la estrategia de clonación se optó por no fusionar las proteasas recombinantes a la cola de histidinas codificada por el vector, ya que el proceso de purificación se preveía fuese sencillo debido al bajo nivel de secreción de proteínas endógenas de *P. pastoris*.

4.4.1 Expresión empleando el vector pPICZ α

El vector pPICZ α fue utilizado para la expresión y secreción de las proteasas aspárticas de *G. verum* y *F. carica* preprogaline A y preprofigain A respectivamente. Este vector permite elevados niveles de expresión del gen de interés mediante la inducción del promotor AOX1 mediante metanol. Los cebadores utilizados para la amplificación del ADNc de *G. verum* fueron pGv1-EcoRI/ pGv1-XbaI y en el caso de *F. carica* pFc1-EcoRI/pFc1-XbaI y Fc1-EcoRI/pFc1-XbaI, estos últimos para amplificar el gen sin el péptido señal ni el prosegmento (Tabla 10). Los productos de PCR fueron digeridos con las enzimas de restricción EcoRI y XbaI y ligados al vector digerido utilizando las mismas enzimas, de forma que en la secuencia de las proteasas recombinantes se introducían varios aminoácidos “extra” entre el factor- α y la secuencia de la proteasa (indicados en rojo en la Figura 39).

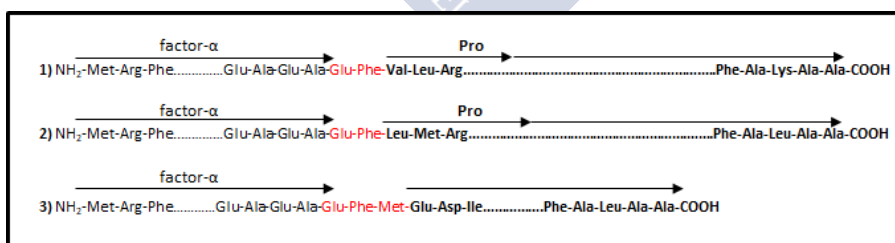


Figura 39. Secuencia de las proteasas recombinantes generadas utilizando el vector pPICZ α . Las proteasas recombinantes 1) progaline A, 2) profigain A y 3) figain A contenían en su secuencia el péptido señal factor- α de *S. cerevisiae*, seguido de varios aminoácidos (en rojo) introducidos debido a la estrategia de clonación.

Una vez comprobado que la secuencia de los plásmidos destinados a la transformación (pPICZ α -progv1AP, pPICZ α -profc1AP y pPICZ α -fc1AP) obtenidos desde *E. coli* era correcta, estos fueron cortados con la enzima de restricción SacI y utilizados para la electroporación de *P. pastoris*.

El ADN genómico de los transformantes fue analizado para verificar la integración de los genes a expresar. La Figura 40 muestra el análisis mediante PCR del ADN genómico de las células de *P. pastoris* transformadas con la construcción pPICZ α -progv1AP, utilizando el cebador específico pMgv1f y el cebador 5'AOX1, que dan lugar a la amplificación de un fragmento de aproximadamente 1500 pb.

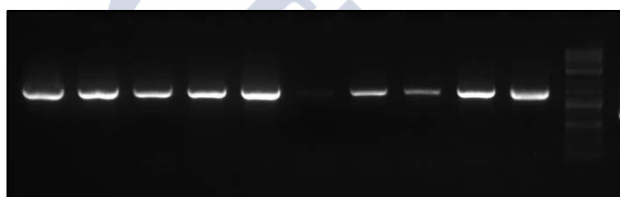


Figura 40. Análisis mediante PCR del ADN genómico de los clones de *P. pastoris* transformada con la construcción pPICZ α -progv1AP.

Tras comprobar la integración de los genes en el genoma de *P. pastoris* se estudió su capacidad para metabolizar el metanol. En el proceso de transformación pueden tener lugar fenómenos de recombinación doble entre las regiones homólogas del vector recombinante y del genoma de la levadura, pudiéndose producir la sustitución (en lugar de la integración) del gen AOX1 silvestre por la construcción génica introducida, y con ello la interrupción del gen que codifica la alcohol oxidasa 1, lo cual daría lugar a la aparición del fenotipo Mut^s (Methanol utilization slow). La transformación con el vector linealizado en la región 5'AOX1 mediante la enzima de restricción SacI favorecerá que se produzca la inserción génica y no el reemplazamiento del gen AOX1, siendo la mayoría de los

transformantes obtenidos Mut⁺ (Methanol utilization plus) (Romanos 1995). El fenotipo de los transformantes se analizó mediante una PCR utilizando el ADN genómico y los cebadores 5'AOX1 y 3'AOX, comprobándose que todos los clones a excepción de uno eran Mut⁺. En la Figura 41 puede apreciarse que en la mayoría de los clones analizados aparecen dos bandas, una correspondiente al gen AOX1 (~2200 pb) y otra correspondiente al gen de la proteasa clonada, con un tamaño de 1979 pb en el caso de la cepa transformada con la construcción pPICZα-progv1AP (A, calles 1-7), 1980 pb en el caso de la transformada con pPICZα-profc1AP (B, calles 1-9) y 1858 pb en el caso de pPICZα-fc1AP (calles 1-5 y calles 7-9); en el clon 6 (calle 6) correspondiente a esta transformación parece que no se ha integrado la construcción y únicamente se aprecia la banda correspondiente al gen AOX1 silvestre. En la amplificación realizada utilizando como molde el genoma de la cepa transformada con el vector pPICZα sin inserto, utilizada como control negativo, se observa una banda muy marcada, de aproximadamente 600 pb, que correspondería a la parte del vector pPICZα que se integra en el genoma y otra mucho más tenue, que corresponde al gen AOX1 silvestre (C, calles C-1 y C-2).

Una vez confirmada la integración de las construcciones en el genoma de *P. pastoris* X33 y deducido el fenotipo Mut mediante PCR, que condiciona las condiciones de inducción del cultivo, se abordó la expresión de las proteasas aspárticas recombinantes, utilizando los medios de cultivo BMGY/BMMY, BMG/BMM y MGY/MM. Los dos primeros tamponados con tampón fosfato pH 6 y el último sin tamponar, para que durante el crecimiento se produjese una bajada de pH. Se añadió metanol 100 % a una concentración final de 0.5 % cada 24 horas para mantener la inducción y se recogieron los sobrenadantes a diferentes tiempos (6 h, 24 h, 36 h, 48 h, 60 h, 72 h y 96 h).

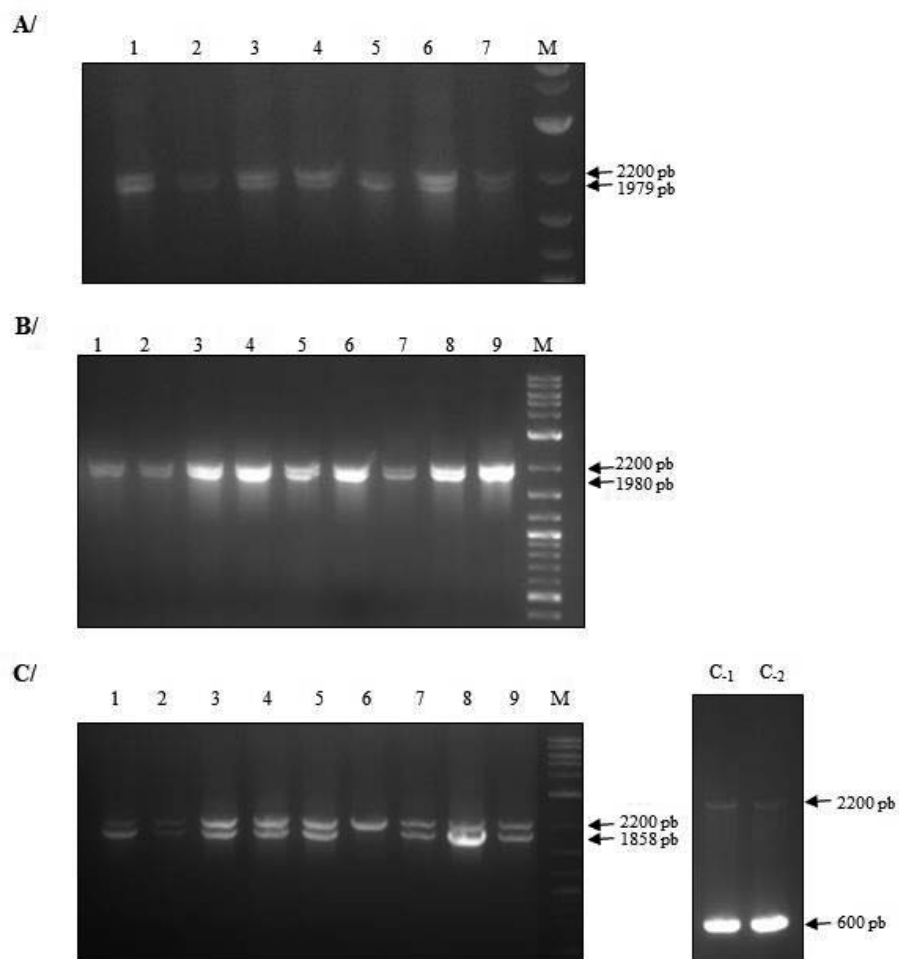


Figura 41. Análisis del fenotipo Mut de las cepas de *P. pastoris* transformadas con las construcciones pPICZα-progv1AP, pPICZα-profc1AP y pPICZα-fc1AP. El análisis se realizó mediante PCR utilizando los cebadores 5' AOX1 y 3' AOX1 y como molde el ADN genómico de *P. pastoris* transformada con las construcciones (A) pPICZα-progv1AP (calles 1-7 corresponden a los clones 1-7), (B) pPICZα-profc1AP (calles 1-9: clones 1-9), (C) pPICZα-fc1AP (calles 1-9: clones 1-9). Las calles de la imagen C, C₁ y C₂, corresponden a la levadura transformada con el vector sin inserto.

Los sobrenadantes de las cepas recombinantes, que deberían contener las proteasas recombinantes, procedentes del medio MM se dializaron y concentraron frente a tampón fosfato potásico a pH 6, mientras que los procedentes de los otros dos medios de cultivo, que estaban tamponados a pH 6, se dializaron previamente frente a tampón citrato a pH 3 y se mantuvieron 2 horas a 37 °C para activarlos, transcurrido este tiempo se siguió el mismo procedimiento de diálisis que para el medio MM. No se detectaron mediante SDS-PAGE bandas que pudiesen corresponder a las proteasas recombinantes, ni en el sobrenadante del medio de cultivo ni en el interior celular (Figura 42). El ensayo de actividad utilizando FTC-k-caseína como sustrato realizado con los sobrenadantes de las cepas transformadas no mostró diferencias con los sobrenadantes de la cepa silvestre.

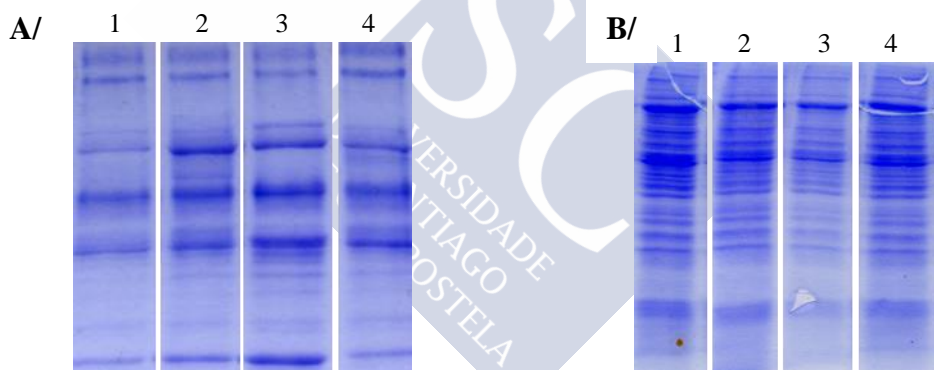


Figura 42. Análisis de la expresión de la preprofigain A, figain A y progaine A en *P. pastoris* utilizando el vector pPICZα. (A) Sobrenadantes de la levadura cultivada durante 60 horas en medio de inducción MM y (B) interior celular de las cepas cultivadas durante 6 horas en medio de inducción MM. Cepa silvestre (calle 1), cepas transformadas con pPICZα-profc1AP (calle 2), con pPICZα-fc1AP (calle 3) y con pPICZα-progv1AP (calle 4).

A pesar de que no se detectaba la expresión de los genes se comprobó que estos se estaban transcribiendo (Figura 43).

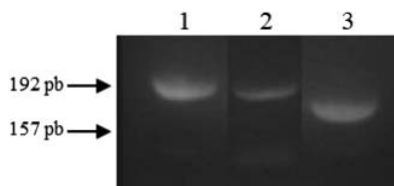


Figura 43. Análisis de la transcripción de los genes *fc1AP* y *gv1AP* de cepas de *P. pastoris* transformadas con las construcciones realizadas con el vector pPICZα. RT-PCR de cepas de *P. pastoris* transformadas con pPICZα-profc1AP (calles 1) y pPICZα-fc1AP (calle 2) utilizando los cebadores pMfc1f/ pMfc1r, y pPICZα-progv1AP (calle 3) utilizando los cebadores pMgv1f/ pMgv1r.

4.4.2 Expresión empleando el vector pGAPZα.

Se abordó la expresión de las proteasas recombinantes con una secuencia idéntica a las obtenidas empleando el vector pPICZα (Figura 39) pero bajo el promotor constitutivo del gen gliceraldeído-3-fosfato deshidrogenasa (PGAP) (Figura 15), esta enzima se expresa de manera constitutiva en *Pichia* cuando crece bajo fuentes de carbono como glicerol o glucosa. Los niveles de expresión alcanzados utilizando glucosa pueden ser mayores a los obtenidos con la enzima alcohol oxidasa cultivada con metanol (Waterham et al. 1997). Las construcciones génicas pGAPZα-progv1AP, pGAPZα-profc1AP y pPGAPZα-fc1AP fueron obtenidas de forma análoga a las del apartado anterior. Con este vector se realizó una cuarta construcción, pGAPZα-progv2AP, para la expresión de la preprogaline B, utilizando los cebadores pGv2-EcoRI/pGv2-XbaI. Esta proteasa recombinante constaría de la señal de secreción factor-α, seguida de los aminoácidos Glu-Phe, codificados por los nucleótidos correspondientes al sitio de restricción EcoRI, utilizado para la clonación, fusionado al fragmento de la preprogaline B que incluye los 7 últimos aminoácidos (Ser-Ala-Ala-Leu-Leu-Ser-Ala) correspondientes al péptido señal nativo hasta el codón stop (Figura 44).

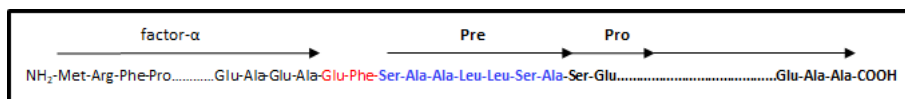


Figura 44. Secuencia de la preprogalina B recombinante generada utilizando el vector pGAPZα. La secuencia contiene el factor-α de *S. cerevisiae*, en rojo aminoácidos introducidos debido al sitio de restricción utilizado para la clonación, en azul los últimos 6 aminoácidos del péptido señal nativo (Pre) seguidos del prosegmento (Pro) y la secuencia de la proteasa madura.

Los insertos amplificados mediante PCR se clonaron en el vector pCR®-Blunt II TOPO®, tras lo cual fueron digeridos con las enzimas de restricción EcoRI y XbaI y ligados al vector pGAPZα digerido con las mismas enzimas. Las construcciones fueron utilizadas para transformar *E. coli* y los plásmidos obtenidos analizados mediante enzimas de restricción y secuenciados antes de proceder a su corte con la enzima BspHI (Figura 45) para su linearización e introducción en *P. pastoris* X-33.

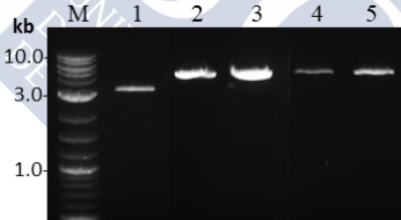


Figura 45. Construcciones derivadas del vector pGAPZα linearizadas con la enzima BspHI. Plásmidos pGAPZα (calle 1), pGAPZα-progv1AP (calle 2), pGAPZα-progv2AP (calle 3), pGAPZα-profc1AP (calle 4) y pGAPZα-fc1AP (calle 5). M: marcador 2-Log DNA Ladder.

La secuenciación de los plásmidos verificó que las construcciones eran las esperadas.

El ADN genómico de las cepas de *P. pastoris* transformadas con las diferentes construcciones fue analizado mediante PCR, confirmándose la integración de estas en el genoma de la levadura (Figura 46).

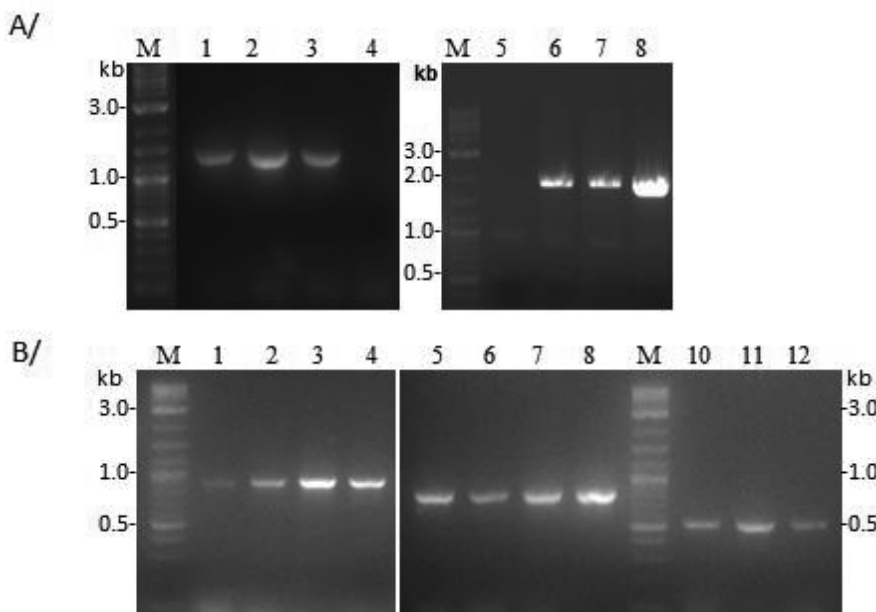


Figura 46. Análisis mediante PCR del ADN genómico de las cepas de *P. pastoris* transformadas con las construcciones (A) pGAPZα-progv1AP (calles 1-4), pGAPZα-progv2AP (calles 5-8) y (B) pGAPZα-profc1AP (calles 1-4) y pPGAPZα-fc1AP (calles 5-8). El ADN genómico de las calles 10-12 corresponde a cepas que fueron transformada con el vector pGAPZα sin inserto. El análisis se realizó utilizando las parejas de cebadores pGAP/pMgv1r en el caso de cepas transformadas con la construcción pGAPZα-progv1AP (1392pb), pGAP/3'AOX1 para pGAPZα-progv2AP (1950pb), pGAP/pMfc1r para pGAPZα-profc1AP (912pb) y pPGAPZα-fc1AP (804pb), y pGAP/3'AOX1 para el vector pGAPZα (540pb).

Las cepas fueron cultivadas en YPD durante 3 días y se analizó la secreción de las proteasas recombinantes al medio de cultivo mediante SDS-PAGE y un ensayo de actividad sobre κ -caseína. El análisis de las proteínas del medio de cultivo (Figura 47) mostró la presencia de un péptido con un tamaño aproximado de 31 kDa (calle 2) en la cepa transformada con la construcción pGAPZα-progv2AP que no estaba presente en la cepa silvestre; en el resto de las cepas recombinantes (calles 3-5) el patrón de bandas era el mismo que el observado en la cepa silvestre (calle 1). Únicamente se detectó

actividad κ -caseinolítica en el sobrenadante de la cepa transformada con la construcción pGAPZ α -progv2AP; esta actividad era inhibida cuando el sobrenadante era preincubado con el inhibidor de proteasas aspárticas pepstatina A; en el resto de las cepas, incluida la cepa silvestre utilizada como control, la actividad proteolítica era prácticamente nula.

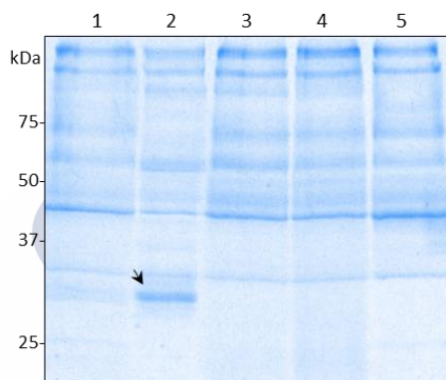


Figura 47. Análisis de la expresión de la progaline B, progaline A, profigain A y figain A en *P. pastoris* utilizando el vector pGAPZ α . SDS-PAGE de las proteínas del medio de cultivo de *P. pastoris* silvestre (calle 1) y transformada con las construcciones pGAPZ α -progv2AP (calles 2), pGAPZ α -progv1AP (calles 3), pGAPZ α -profc1AP (calle 4) y pPGAPZ α -fc1AP (calle 5) incubada durante 3 días a 28 °C y 230 rpm. La flecha indica la banda que aparece en la cepa transformada con pGAPZ α -progv2AP, pero no en la silvestre. El sobrenadante del medio de cultivo (1 mL) fue precipitado mediante TCA y analizado en condiciones reductoras. En el margen izquierdo de la imagen figuran los pesos moleculares del marcador.

La ausencia de actividad proteolítica en los sobrenadantes de las cepas transformadas con las construcciones pGAPZ α -progv1AP, pGAPZ α -profc1AP y pPGAPZ α -fc1AP podía ser debida a que los genes no se estaban expresando, o bien que lo hiciesen con baja eficiencia y las proteasas estuviesen en su forma inactiva. Para comprobar esta segunda hipótesis se acidificaron los sobrenadantes mediante adición de ácido clorhídrico hasta alcanzar pH 3 y se incubaron a 37 °C durante 12 horas, transcurridas las cuales los

sobrenadantes se neutralizaron a pH 6. La actividad κ -caseinolítica seguía siendo nula tras el proceso de activación.

Se comprobó que los genes *progv1AP* y *profc1AP* se estaban transcribiendo (Figura 48), por lo que el problema en la expresión debía estar asociado al proceso de traducción, o bien en la secreción de las proteasas recombinantes al medio de cultivo. La comprobación de la transcripción del producto de la integración de *pGAPZ α -fc1AP* no fue realizada.

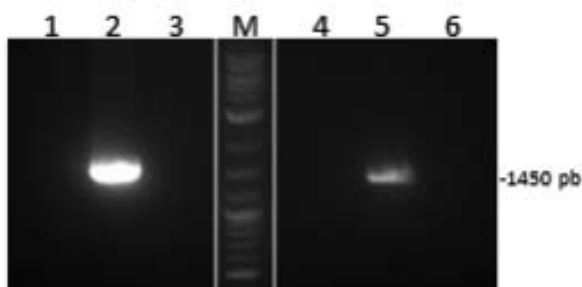


Figura 48. RT-PCR de las cepas transformadas con las construcciones *pGAPZ α -progv1AP* (calle 2) y *pGAPZ α -profc1AP* (calle 5). Las parejas de cebadores utilizados fueron *pGv1-EcoRI/pGv1-XbaI* y *pFc1-EcoRI/pFc1-Xba1* respectivamente. En ambos casos se realizaron controles negativos utilizando ARN al que no se le añadió transcriptasa inversa (calles 1 y 4) y utilizando ADNc obtenido de la cepa silvestre y los cebadores *pGv1-EcoRI/pGv1-XbaI* (calle 3) y *pFc1-EcoRI/pFc1-Xba1* (calle 6). En el margen derecho se indica el tamaño aproximado de los productos de PCR, que se corresponde con el esperado para la amplificación de ambos genes.

4.5 Producción y caracterización de la progaline B en *P. pastoris*

Los sobrenadantes de cuatro clones de *P. pastoris* transformados con *pGAPZ α -progv2AP* fueron analizados mediante SDS-PAGE (Figura 49) y se realizaron ensayos de actividad enzimática a lo largo de tres días de crecimiento. Transcurrido un día de cultivo se detectaban dos péptidos, de 37 y 30.7 kDa, en los sobrenadantes obtenidos desde los clones 3, 5 y 7 (calles 3-5) que no estaban presentes en la cepa silvestre (calle1). El péptido de 30.7 kDa

era más evidente con el transcurso del tiempo (calles 3-5 a las 48 horas y 72 horas de cultivo), mientras que el de 37 kDa únicamente se detectaba en determinados momentos durante el crecimiento de la levadura. El clon1 (calles 2, 7 y 12) presentaba el mismo patrón de bandas que la cepa silvestre (calles 1, 6 y 11), indicando que no estaba secretando la progaline B.

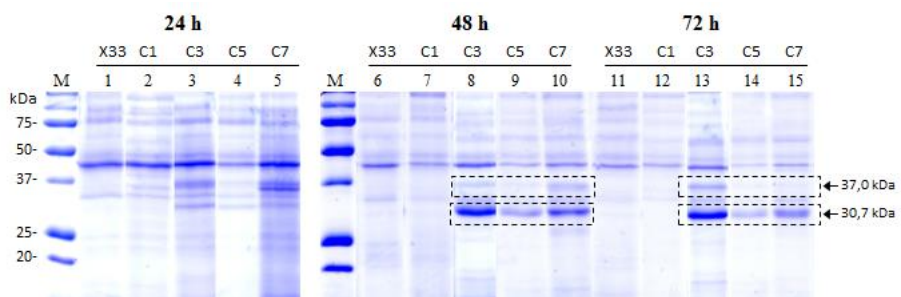


Figura 49. Análisis mediante SDS-PAGE de los clones de *P. pastoris* transformados con la construcción pGAPZα-progv2AP. Los sobrenadantes (2mL) fueron obtenidos mediante centrifugación a las 24 horas (calles 1-5), 48 horas (calles 6-10) y 72 horas (calles 11-15) de incubación y precipitados mediante TCA. Las calles 1, 6 y 11 corresponden al sobrenadante de la cepa silvestre (X33); las calles 2, 7 y 12 al clon1 (C1), calles 3, 8 y 13 al clon 3 (C3), calles 4, 9 y 14 al clon 5 (C5) y calles 5, 10 y 15 al clon 7 (C7). M indica el marcador de pesos moleculares. Las cajas con líneas discontinuas encierran las bandas que aparecen en los clones recombinantes y no en la cepa silvestre.

Los sobrenadantes de las cepas cultivadas durante 3 días que presentaban un patrón de bandas (Figura 49) diferente a la cepa silvestre, mostraron actividad coagulante de la leche transcurridas aproximadamente 5 horas de incubación a 37 °C (Figura 50, tubos T3, T4 y T5), mientras que la leche incubada con el sobrenadante de la cepa silvestre (T6) o con los sobrenadantes de los clones 1 y 2 (T1 y T2) permaneció líquida aun cuando habían transcurrido 15 horas de incubación. Los sobrenadantes producían la coagulación de la leche sin necesidad de realizar un proceso previo de activación mediante bajada de pH. La preincubación de las muestras con el inhibidor de

proteasas aspárticas pepstatina impedía la coagulación (las imágenes no se muestran), lo cual ratifica que esta se produce por acción de una proteasa de tipo aspártico. Se realizó la misma prueba, pero con un paso de activación previa de los sobrenadantes (bajada de pH a 3.5 con HCl, incubación durante 3.5 horas a 37 °C y ajuste a pH 6 con NaOH). El resultado obtenido fue el mismo que en la prueba anterior, siendo improbable la posibilidad de que los clones 1 y 2 estuviesen produciendo la proteasa en su forma inactiva.

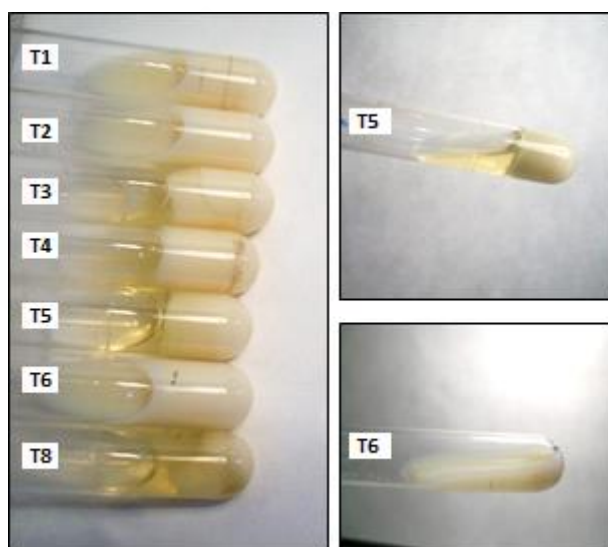


Figura 50. Actividad coagulante de la leche de los sobrenadantes de *P. pastoris* transformada con el gen de la progaine B. La actividad fue ensayada utilizando 24 % leche desnatada pH 6 suplementada con CaCl_2 20 mM a la que se le añadió un volumen de los sobrenadantes obtenidos desde las cepas transformadas mediante centrifugación (clones 1, 2, 3, 5, 7 corresponden a T1-5 de la imagen izq.) y se incubó a 37 °C. El tubo T6 corresponde al sobrenadante de la cepa silvestre (control negativo) y el tubo T8 al sobrenadante de una cepa transformada con el gen de la quimosina de búfalo cultivada en las mismas condiciones que las anteriores (control positivo).

El clon destinado a la producción y caracterización de la progaine B recombinante se seleccionó en base a la actividad del

sobrenadante mediante el ensayo de la FTC- κ -caseína. Se realizaron cultivos por triplicado de los clones 3, 5 y 7 inoculados a la misma densidad óptica inicial y se tomaron muestras a diferentes tiempos. El clon 7 fue seleccionado para los experimentos posteriores, ya que, a pesar de presentar un crecimiento más lento, medido mediante densidad óptica, la actividad κ -caseinolítica transcurridas 72 horas de cultivo era mayor que en los otros clones analizados (Tabla 15).

Tabla 15. Ensayo FTC- κ -caseína* de los sobrenadantes de *P. pastoris* expresando la progaline B.

	DIA 1	DIA2	DIA3	DIA6
X33	26 \pm 5	-5 \pm 5.6	8 \pm 5	58 \pm 1
Clon 3	128 \pm 19	185 \pm 24	162 \pm 17	55 \pm 34
Clon 5	85 \pm 5	170 \pm 28	133 \pm 22	93 \pm 25
Clon 7	83 \pm 20	299 \pm 25	552 \pm 23	204 \pm 7

* La incubación se realizó durante 2 horas a 37 °C, como blancos se utilizaron muestras a las que se añadió el sobrenadante del medio de cultivo en el momento de detener la reacción con TCA.

Un zimograma realizado utilizando el sobrenadante concentrado (15x) del clon 7 transcurridas 72 horas de cultivo, mostraba la presencia de una única banda con actividad proteolítica de un peso molecular aproximado de 50 kDa (Figura 51).

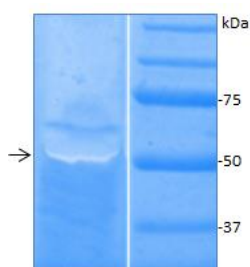


Figura 51. Zimograma del sobrenadante de cultivo obtenido desde el clon 7 de *P. pastoris* expresando la progaline B. La banda con actividad proteolítica se indica con una flecha.

El clon 7 de la cepa de *P. pastoris* productora de la progaline B se ha depositado según el Tratado de Budapest sobre el reconocimiento internacional del depósito de microorganismos con fines del procedimiento en materia de patentes en la CECT (Colección

Española de Cultivos Tipo) bajo la referencia de identificación PROGVB-7 y con el número de orden CECT 13141.

4.5.1 Estructura

El sobrenadante de la cepa recombinante PROGVB-7, colectado transcurridos tres días de cultivo, fue analizado mediante Western blot utilizando suero de conejo conteniendo anticuerpos policlonales contra el péptido sintético PV13, que presenta la secuencia codificada por los aminoácidos 243 a 255 de la preprogaline B (la numeración hace referencia a la secuencia de la proteasa nativa) (Figura 52).

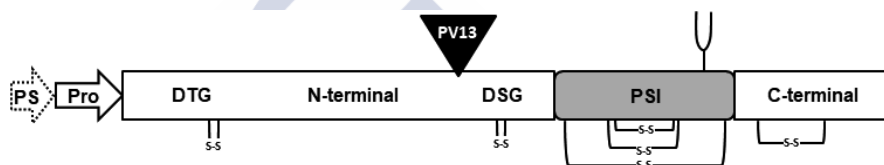


Figura 52. Localización de la secuencia de la preprogaline B utilizada para la producción de anticuerpos empleados en el Western Blot. El triángulo negro indica la posición del péptido utilizado para generar el anticuerpo PV13. La horquilla indica un sitio de N-glicosilación.

El Western blot realizado al sobrenadante bajo condiciones reductoras, detectaba dos polipéptidos, uno de 37 kDa y otro de 30.7 kDa (Figura 53, calle 1), lo que confirma los resultados obtenidos cuando se comparó el patrón de bandas de la cepa silvestre y las recombinantes (Figura 49) mediante SDS-PAGE. Bajo condiciones no reductoras se detectaba una banda heterogénea con dos polipéptidos de 48 y 43 kDa (Figura 53, calle 2). Estos resultados sugerían que la progaline B está compuesta por dos cadenas, una cadena pesada (HC), con un tamaño molecular de 30.7 kDa, y una cadena ligera (LC) unidas mediante puentes disulfuro. El polipéptido de 37 kDa podría corresponder a la HC conteniendo el prosegmento; los polipéptidos de 48 y 43 kDa representarían respectivamente las formas glicosilada y

no glicosilada de la progline B con las cadenas HC y LC unidas mediante puentes disulfuro.

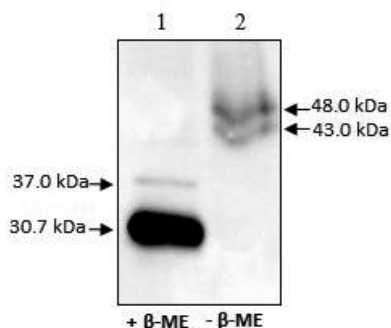


Figura 53. Análisis mediante Western blot del sobrenadante de la cepa de *P. pastoris* PROGVB-7 transcurridas 72 h de cultivo. El sobrenadante fue concentrado mediante precipitación con TCA. Las proteínas fueron analizadas en presencia (+β-ME, calle1) y ausencia (-β-ME, calle2) de β-mercaptoetanol.

Para confirmar la interpretación de la estructura de la progline B realizada en base al Western blot se llevaron a cabo dos ensayos: (i) una espectrometría de masas MALDI-TOF y (ii) la deglicosilación de la proteasa recombinante.

(i) Espectrometría de masas: Las bandas correspondientes a los polipéptidos de 37 y 30.7 kDa fueron escindidas desde el SDS-PAGE, digeridas con tripsina y analizadas mediante MALDI-TOF. Las relaciones masa/carga (m/z) de los péptidos tripticos determinadas experimentalmente se compararon con las esperadas a partir de una digestión teórica realizada *in silico* utilizando la secuencia de aminoácidos de la progline B, y teniendo en cuenta la especificidad de la tripsina, que hidroliza el enlace peptídico por el extremo C-terminal de la lisina (K) y la arginina (R), excepto cuando el residuo siguiente es una prolina. Los análisis revelaron que el polipéptido de 37 kDa contenía varios péptidos tripticos incluidos en la secuencia de aminoácidos AKLDQVNRLAGSVPR, que no estaban presentes en las muestras procedentes de la digestión del polipéptido de 30.7 kDa, el resto de los péptidos detectados aparecían en las muestras extraídas desde ambas bandas del gel (Tabla 16).

Tabla 16. Análisis mediante MALDI-TOF de los polipéptidos de 37 kDa y 30.7 kDa de la progaline B recombinante. Relación m/z de los péptidos tripticos (teóricos y experimentales) y su secuencia de aminoácidos. Las secuencias subrayadas en gris fueron detectadas únicamente en la banda de 37 kDa.

m/z teórica	m/z experimental		Intensidad		**Residuos	P*	Secuencia de la progaline B
	Banda 37kDa	Banda 30.7kDa	Banda 37kDa	Banda 30.7kDa			
3068.43	3069.43	3069.41	73.23	26.257	300-327	1	HTDEEQGGLVFGGVDPSPHFKEHVYAK
2284.04	2285.03	2285.03	3322.30	2707.14	300-320	0	HTDEEQGGLVFGGVDPSPHFK
2200.04	2201.03	2201.04	160.72	28.04	212-232	0	SAEIHVGTGSIAGFFSQDSVK
2107.06	2108.06	2108.06	1833.20	2161.80	240-258	1	DQDFEATREPGITFVAAK
1817.95	1818.95	1818.95	114.64	124.03	233-248	1	LGDLVVKDQDFEATR
1622.92	1623.92	-----	282.20	-----	111-125	2	AKLDQVNRLAGSVPR
1423.79	1424.78	-----	150.10	-----	113-125	1	LDQVNRLAGSVPR
1205.63	1206.62	1206.62	1248.61	523.13	275-284	0	AVPVWYNMVK
1093.50	1094.60	1094.51	8586.36	3722.89	240-248	0	DQDFEATR
1031.56	1032.57	1032.55	98.37	48.67	249-258	0	EPGITFVAAK
942.52	943.53	-----	12709.31	-----	111-118	1	AKLDQVNR
866.45	867.45	867.45	1991.42	395.62	392-399	0	SVSQYVK
802.40	803.40	803.38	211.83	104.22	321-327	0	GEHVYAK
743.45	744.40	-----	1919.98	-----	113-118	0	LDQVNR

P*: puntos de corte de la tripsina no efectuados permitidos
 **Residuos: La posición de los residuos puede consultarse en la Figura 54.

Esta secuencia de aminoácidos, presente en el polipéptido de 37 kDa y no en el de 30.7 kDa, forma parte del prosegmento (Figura 54). En el análisis también se detectaba una secuencia de aminoácidos perteneciente al inicio del PSI en ambas bandas.

```

1  MRFFSIFTAV LFAASSALAA PVNTTTEDET AQIPAEAVIG YSDLEGDFDV AVLPFSNSTN NGLLFINTTI
71  ASTAAKEEGV SLEKREAEAE FSAALLSASE TKLVRVSLKK AKLDQVRLA GSVPRGSARI NRLYGNMGSS
141  DSDIIALKNY MDAQYYGEIG IGTPAQKFIV IFDTGSSNLW VPSTKCYFSP ACYFHSKYNS KHSSTYKNG
211  KSAEIHYGIG SIAGFFSQDS VKLGDVVKD QDFIEATREP GITFVAAKFD GLLGLGFQEI SVGKAVFVWY
281  NMVQGLVAE PVFSFWENRH IDEEQGGLV FGGVDPSEFK GEHVYAKVTK KGMWQDFMDG VLVGGESTGF
351  CANGCSAIAD SGTSLLAGPT AVVVQINHAI GATCVVSQEC KSVVSQYKGI ILEMLIAKTE KTKICASIGL
421  CLPGRSRDDI MIIESVVEET DSASDGLRND EYCSAQMAV VMQONQVRQN QTEEKILDYI NQLCDRLPSE
491  MGESAVDCSS LASLPKIAFS IGGKAFELSP EYVLKVGEG AAAQCISGFT ALDVAPPLGP LNIILGVTFM
561  SYHIVFDYGN MRVGFEEAA

```

Figura 54. Secuencia de aminoácidos de la progline B recombinante mostrando los péptidos tripticos identificados mediante MALDI-TOF (Tabla 16). Los péptidos identificados en el fragmento de 37 kDa se muestran subrayados en rojo, y en verde los correspondientes al fragmento de 30.7 kDa. Los residuos pertenecientes al factor- α aparecen en un recuadro, los pertenecientes al prosegmento sobre fondo azul, y los que forman parte del PSI sobre fondo naranja. En el margen izquierdo los números indican la posición del primer residuo en la secuencia.

Los resultados obtenidos mediante MALDI-TOF fueron confirmados mediante espectrometría de masas en tándem (MS/MS) (Roepstorff 1994). La comparación de la fragmentación de los péptidos obtenidos desde ambas bandas mostraba que la diferencia entre ellas se encontraba en la ausencia, en el polipéptido de 30.7 kDa, de las secuencias aminoacídicas que figuran en la Tabla 17, y que pertenecen al prosegmento.

Tabla 17. Péptidos de la progaline B recombinante identificados mediante nano LC-MS/MS presentes en el polipéptido de 37 kDa y ausentes en el de 30.7 kDa.

M/z exp.	Mr exp.	Mr teórica	z	Score	Residuos	P	Secuencia
472.25	942.49	942.52	2+	259	111-118	1	AKLDQVNR
842.43	1682.85	1682.82	2+	16461	133-148	0	LYGNMGSSSDIIALK
689.67	2065.99	2066.05	3+	1082	130-148	1	INRLYGNMGSSSDIIALK

Mr: masa molecular relativa

Score: probabilidad estadística que confirma la secuencia

En la Figura 55 puede consultarse el espectro del fragmento correspondiente a la secuencia INRLYGNMGSSSDIIALK.

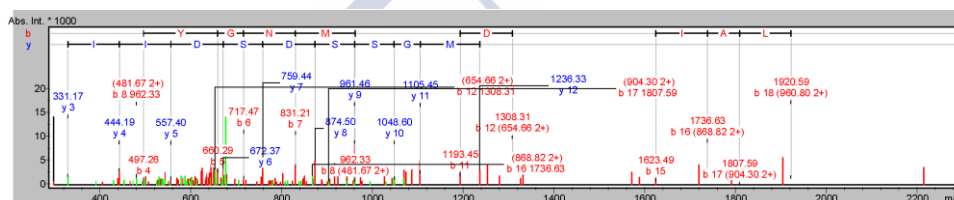


Figura 55. Identificación de la secuencia INRLYGNMGSSSDIIALK mediante espectrometría de masas en tándem (MS/MS) a partir del ion con una relación m/z 689.67. En la parte superior del espectro se muestra la secuencia deducida realizada a partir de la identificación de los iones “b” (rojo, fragmentos NH₂-terminales) y las de iones “y” (azul, fragmentos COOH-terminales).

(ii) Deglicosilación: La progaline B contiene en su secuencia un posible sitio de glicosilación, situado en la parte final del inserto específico de plantas (PSI) (Figura 52). El análisis mediante Western Blot (Figura 56) muestra que la deglicosilación de la proteasa recombinante con la enzima endoglicosidasa H, bajo condiciones no reductoras, da lugar a la desaparición de la banda de 48 kDa, observándose únicamente una banda de 43 kDa (calles 3 y 4), que correspondería a la forma no glicosilada de la progaline B madura. Bajo condiciones reductoras únicamente se observaba un polipéptido de 30.7 kDa, independientemente de que la muestra fuese o no tratada con endoglicosidasa H (calles 1 y 2). Esto es debido a que la reducción de

el/los puente/s disulfuro que mantienen unidas la cadena pesada (HC) y la cadena ligera (LC) resulta en la separación de ambas cadenas, observándose en el Western blot únicamente la HC; el sitio de glicosilación de la progaline B se encuentra localizado en la LC, mientras que los anticuerpos generados frente al péptido PV13 únicamente reconocen la HC, por ello el tamaño de esta es el mismo en la muestra tratada con o sin endoglicosidasa H (calles 1 y 2 respectivamente). En el Western blot también se observaba una banda tenue de 27 kDa, quizás producto de la degradación de la progaline B. Este polipéptido solo era detectado cuando las muestras eran incubadas con endoglicosidasa H (calles 1 y 3), probablemente debido a que la glicosilación protege a la proteína de la degradación proteolítica (Goettig 2016).

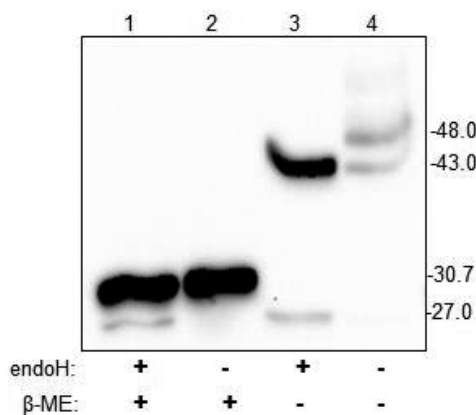


Figura 56. Análisis de N-glicosilación de la progaline B. Western blot de la progaline B incubada con (+) o sin (-) endoglicosidasa H y en presencia (+) o ausencia (-) de β-mercaptoetanol. Los números en el margen derecho indican los pesos moleculares estimados de los péptidos.

La purificación de la progaline B desde el sobrenadante del medio de cultivo se llevó a cabo mediante dos pasos. (i) En primer lugar se empleó una columna de intercambio aniónico cargada con DEAE Bio-Gel A (Bio-Rad), aquellas fracciones con actividad proteolítica fueron analizadas mediante SDS-PAGE y se seleccionaron las que presentaban

un menor número de proteínas contaminantes. (ii) Estas fracciones fueron aplicadas a una columna monoQ (HR 5/5 Amersham Bioscience) utilizando un sistema FPLC. Las proteínas eluidas fueron dializadas frente a tampón fosfato sódico 20 mM pH 7.5 y su pureza se comprobó sometiéndolas a electroforesis en un gel SDS-PAGE 15 % en condiciones reductoras y tinción con plata (Figura 57). Esta técnica nos permitió estimar el peso molecular de la cadena ligera (13.5 kDa), la cual no era detectada utilizando tinción con azul de Coomassie.

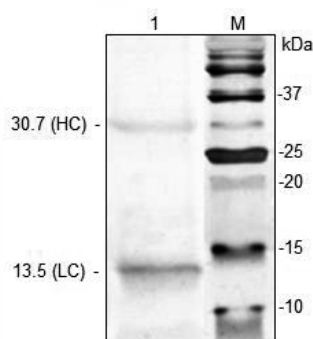


Figura 57. Análisis mediante SDS-PAGE de la progaline B purificada. La proteasa fue visualizada utilizando tinción con plata, detectándose los dos polipéptidos de 30.7 kDa (HC) y 13.5 kDa (LC). Los números en el margen izquierdo de la imagen indican el peso molecular de las proteínas del marcador (M).

4.5.2 Producción en medio de cultivo YPD

La producción de la progaline B en medio YPD fue analizada a lo largo de las diferentes fases de crecimiento de la levadura y se estudió el efecto de los cambios de pH del medio de cultivo sobre la estructura y actividad de la enzima (Figura 58.A). La progaline B secretada al medio alcanzaba su máxima actividad cuando habían transcurrido entre 70 y 92 horas de incubación de la levadura a 28 °C y 230 rpm, durante la fase estacionaria de crecimiento, después la actividad decrecía progresivamente. El pH inicial del medio YPD fue de 6.4, y decrecía hasta alcanzar un mínimo de 4.6 transcurridas 27 horas de crecimiento, luego se iba incrementando gradualmente con el crecimiento de la

levadura, hasta alcanzar un valor de 7.58 transcurridas 195 horas de incubación. Para establecer la relación entre el procesamiento de la progaline B, los cambios de pH del medio y la actividad de la proteasa se realizó un análisis mediante Western blot al sobrenadante del medio de cultivo en diferentes momentos del crecimiento de la levadura recombinante (Figura 58.B). El Western blot detectaba un polipéptido de 30.7 kDa transcurridas 48 horas de cultivo, el cual corresponde a la HC de la enzima totalmente procesada, sin el factor- α , ni el péptido señal, y con la escisión parcial del PSI. Cuando la muestra se analizó bajo condiciones no reductoras se apreciaban dos bandas, correspondientes a las formas glicosilada y no glicosilada de la HC unida a la LC mediante puentes disulfuro. El incremento de pH en el medio daba lugar a la acumulación de un polipéptido de 37 kDa, que corresponde a la HC con el prosegmento. Transcurridas 195 horas de incubación, se apreciaba una banda adicional en la muestra analizada bajo condiciones no reductoras, que correspondería a la progaline B no procesada.

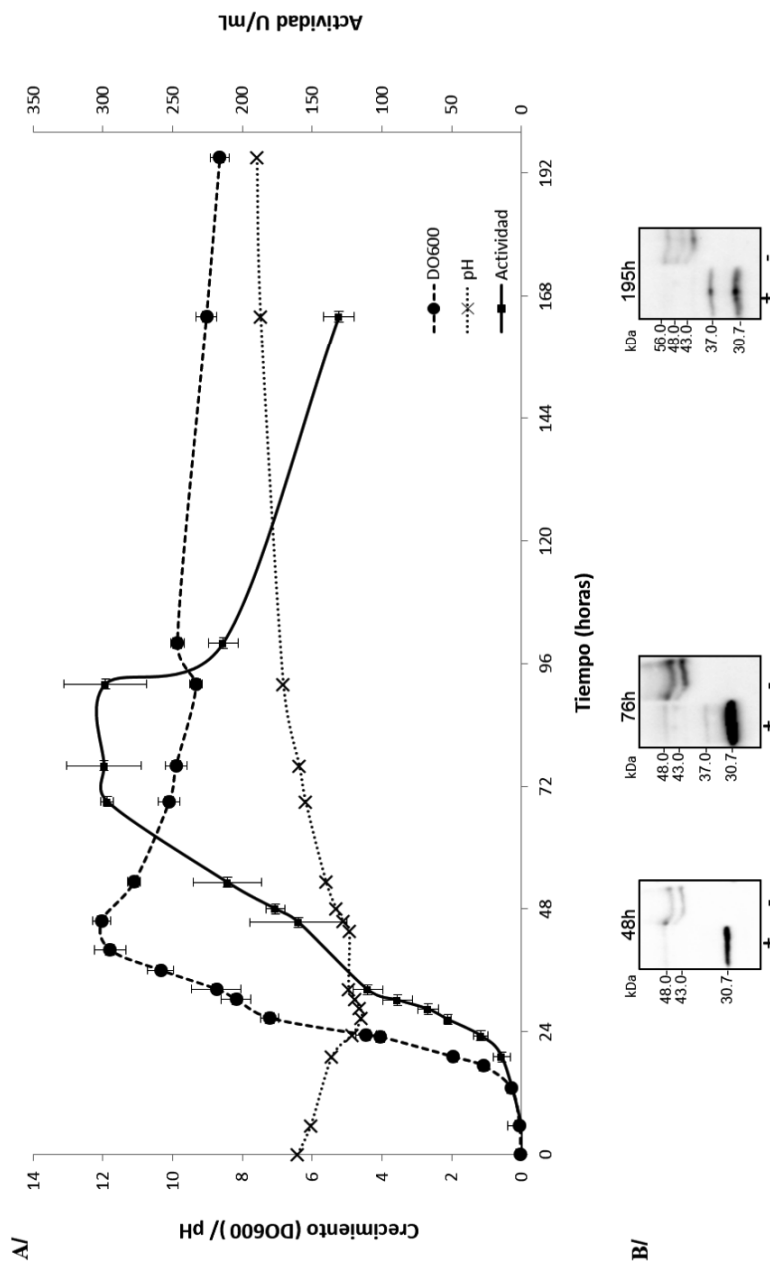


Figura 58. Curva de crecimiento de la levadura recombinante y análisis de la progaline B secretada. (A) Crecimiento de *P. pastoris* (DO₆₀₀, ---●---) en medio YPD, actividad K-caseinólítica (—●—) de la progaline B secretada y pH del medio de cultivo (---X---). Los datos representan la media de tres experimentos independientes. **(B)** Western blot de la progaline B en diferentes momentos del crecimiento. Las muestras (20µL de medio de cultivo) fueron analizadas utilizando el suero generado contra el péptido PVI3. Los signos (+) y (-) indican la presencia o ausencia de β-ME en el tampón de carga del SDS-PAGE. En el margen izquierdo se indica el peso molecular de los péptidos detectados.

4.5.3 Efecto del pH del medio de cultivo en la estructura y actividad

Para comprobar cómo afectaba el pH del medio de cultivo al procesamiento de la progaline B, la levadura se cultivó durante tres días en medio YPD tamponado a pH 7 (YPD7) o a pH 4 (YPD4). El pH de ambos medios fue medido regularmente durante el ensayo, comprobándose que los tampones utilizados en ambos casos, fosfato sódico 100 mM en el caso del medio YPD7 y tampón citrato 100 mM en el caso de YPD4, eran capaces de mantener el pH durante el crecimiento dentro de los valores 6.9–7.2 en el caso del YPD7 o entre 4.0–4.1 en el caso de YPD4.

Cuando el medio de cultivo utilizado fue YPD o YPD4 la progaline B se obtenía desde el medio de cultivo en su forma madura, así, en la Figura 59.A1 (calles 1, 2, 5, 6) podemos observar el patrón de bandas característico de la enzima procesada. La forma no glicosilada de la progaline B detectada en el medio YPD4 (calle 2) aparecía en una cantidad muy baja. Cuando el medio utilizado fue YPD7 (calles 3 y 4), el análisis mediante Western blot bajo condiciones reductoras (calle 3) mostraba dos polipéptidos, de 51 y 56 kDa, cuyos pesos moleculares corresponden al tamaño estimado del precursor de la progaline B (conteniendo el prosegmento y el PSI) sin o con los carbohidratos correspondientes a la N-glicosilación respectivamente. Al igual que en el caso de la enzima detectada en el medio YPD4, la forma no glicosilada (polipéptido de 51 kDa) aparecía en cantidad muy inferior a la glicosilada (56 kDa) (Figura 60). Bajo condiciones no reductoras se detectaba una banda heterogénea de ~58 kDa (calle 4), el incremento en el peso molecular con respecto a los polipéptidos detectados bajo condiciones reductoras podría ser debido a que los puentes disulfuro se mantienen intactos, alterando la migración de la proteína.

La progaline B recombinante secretada por las células de *P. pastoris* cultivadas en YPD o en YPD4 presentaban actividad κ -caseinolítica, mientras que, la progaline B no procesada, obtenida a partir de la cepa recombinante cultivada en el medio tamponado a pH 7,

prácticamente no tenía actividad proteolítica (Figura 59.B, S y SA1). Para comprobar si era posible activarla mediante acidificación, el pH de los sobrenadantes se bajó mediante la adición de HCl hasta un valor de 3.5 y se incubaron durante 4.5 horas a 37 °C antes de realizar el ensayo de actividad. La progaine B que se encontraba en el medio YPD7 fue procesada (en la Figura 61 A.2 calle 3, aparece la banda de 30.7 kDa correspondiente a la HC sin prosegmento) y pasaba a tener actividad proteolítica (177 ± 42) (Figura 59.B SA2). El patrón de bandas del Western blot y la actividad κ -caseinolítica de los sobrenadantes congelados inmediatamente después del cultivo (S), y de aquellos que se mantuvieron 4.5 horas a 37 °C sin ningún tratamiento (SA1) era muy similar.

El crecimiento en YPD7 provocaba la aparición de un polipéptido de 27 kDa, tanto en las muestras sometidas a la acidificación como aquellas no tratadas (en la Figura 61 únicamente se aprecia en A2, calle 3), probablemente procedente de la degradación proteolítica de la progaine B mediante proteasas extracelulares (Ahmad et al. 2014).

La actividad κ -caseinolítica de las proteasas secretadas por la cepa silvestre (Figura 59.B) era prácticamente nula bajo las condiciones ensayadas.

Los resultados obtenidos demuestran que *P. pastoris* cultivada en medio tamponado a pH 7 secreta la forma inactiva de la progaine B, que puede ser activada *in vitro* mediante su incubación a pH ácido.

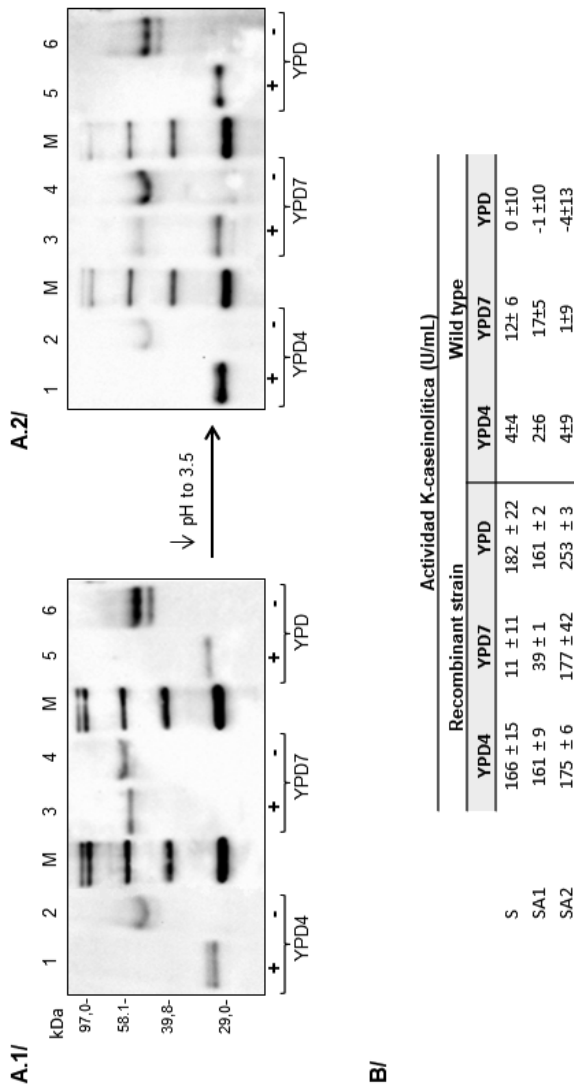


Figura 59. Análisis de la influencia del pH del medio de cultivo en el procesamiento de la progeline B secretada por P. pastoris. (A) Análisis mediante Western blot del medio de cultivo de la levadura recombinante transcurridos 3 días de cultivo en YPD4 (YPD tamponado a pH4), YPD7 (YPD tamponado a pH7) o YPD. Los sobrenadantes de los medios de cultivo fueron incubados a 37 °C durante 4.5 hora, sin tratamiento (A1), o después de ajustar su pH a 3.5 (A2). Las muestras fueron analizadas en presencia (+) o ausencia (-) del agente reductor β -ME. Los números en el margen izquierdo de la imagen A1 corresponden al tamaño de los polipéptidos del marcador de peso molecular (M). (B) Actividad κ -caseinólitica del medio de cultivo de la cepa recombinante o de la cepa silvestre. La línea etiquetada como S corresponde a la actividad de los sobrenadantes y las líneas SA1 y SA2 corresponden a los sobrenadantes analizados en A1 y A2 respectivamente. Los valores fueron obtenidos a partir de dos experimentos independientes realizados por triplicado.

El pH del medio de cultivo influía en la tasa de crecimiento de la levadura recombinante, los cultivos cuyo medio de crecimiento estaba tamponado a pH 4 presentaban una densidad óptica a 600 nm ligeramente superior (11.1 ± 0.4) a aquellos crecidos a pH neutro (8.7 ± 0.5) o en YPD sin tamponar (10.1 ± 0.4) transcurridos 3 días de cultivo.

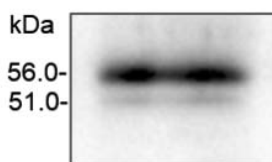


Figura 60. Imagen ampliada del análisis mediante western blot de la progline B obtenida desde el sobrenadante de un cultivo de *P. pastoris* en YPD7 bajo condiciones reductoras.

4.5.4 Determinación del pH de activación

Para determinar el pH de activación de la progline B se utilizó la enzima no procesada obtenida desde el sobrenadante de la cepa recombinante cultivada en medio YPD7 durante 3 días (Figura 59.A1, calles 3 y 4). El sobrenadante de cultivo fue incubado a 28 °C (temperatura de crecimiento de la cepa) en tampones cuyo pH se encontraban en el rango de 3.5-7.0; alícuotas de estas soluciones fueron analizadas transcurridos 30, 150 y 270 minutos (Figura 61.A1, A2 y A3 respectivamente); el sobrenadante incubado a 28 °C durante el mismo tiempo que las muestras fue utilizado como control (A.1, A.2 y A.3, calle 1). El análisis mediante Western blot de los sobrenadantes incubados durante 30 min a un pH comprendido entre 3.5-5.8 revelaba la presencia de un polipéptido de 37 kDa correspondiente a la HC conteniendo el prosegmento (A.1, calles 2-6); cuando el pH de incubación era de 6.6 o 7 únicamente se observaba la

banda correspondiente a la progaline B no procesada (56 kDa) (A.1, calles 6 y 8). Periodos más prolongados de incubación provocaban la escisión del prosegmento, y con ello la aparición del polipéptido de 30.7 kDa (A.2 y A.3). El procesamiento era más lento en los valores de pH más alcalinos, así, transcurridos 150 minutos, la progaline B había sido totalmente procesada a pH 3.5 (A.2, calle 2) y parcialmente en el rango de pH 3.5-5.0 (A.2, calles 3-5), mientras que en pH superiores (5.8 a 7) únicamente se detectaba la forma no procesada (A.2, calles 6-8). Transcurridos 270 minutos de incubación el procesamiento únicamente era completo cuando la enzima se incubaba a pH 3.5 (A.3, calle 2), mientras que la incubación a pH entre 4.3-5.8 daba lugar a la obtención de la enzima en su forma madura y en su forma parcialmente procesada (manteniendo el prosegmento en su estructura) (A.3, calles 3-6), a pH 6.6 y 7 la enzima se mantenía en forma de zimógeno (A.3, calles 7 y 8). La incubación de la progaline B con el inhibidor de proteasas aspárticas pepstatina A prevenía el procesamiento, indicando que la proteasa probablemente sufre un proceso autoproteolítico (A.3, calle 9). Estos resultados indican que la maduración *in vitro* de la progaline B es secuencial, empieza con la rotura parcial del inserto específico de plantas (PSI) y continúa con la escisión del prosegmento.

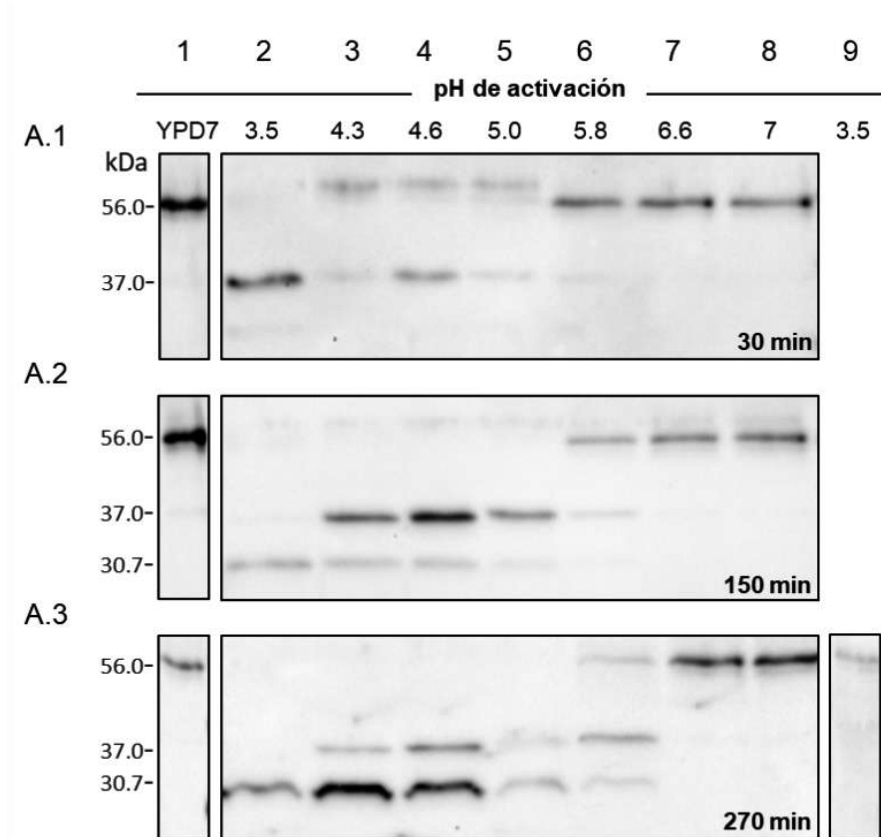


Figura 61. Análisis mediante Western Blot del procesamiento del precursor de la progaline B mediado por pHs. La cepa de *P. pastoris* recombinante fue cultivada en YPD tamponado a pH 7 (YPD7) a 28 °C y 230 rpm durante 3 días, el sobrenadante del cultivo fue ajustado a pH 3.5 (calle 2), 4.3 (calle 3), 4.6 (calle 4), 5.0 (calle 5), 5.8 (calle 6), 6.6 (calle 7), o 7.0 (calle 8), e incubado a 28 °C durante 30 (A.1), 150 (A.2), o 270 minutos (A.3). El sobrenadante incubado a 28 °C fue utilizado como control (calle 1). La muestra de la calle 9, presente únicamente en A.3, fue tratada con pepstatina A antes de ser incubada a pH 3.5.

La eliminación parcial del PSI y la pérdida del prosegmento son necesarios para que el precursor inactivo se convierta en la enzima activa, así a pH entre 6.6 y 7.0, en los cuales la progaline B se encuentra en forma monomérica (A.3, calles 7 y 8) prácticamente no

se detectó actividad enzimática utilizando κ -caseína como sustrato (Figura 62). A medida que el prosegmento era procesado la actividad enzimática se incrementaba.

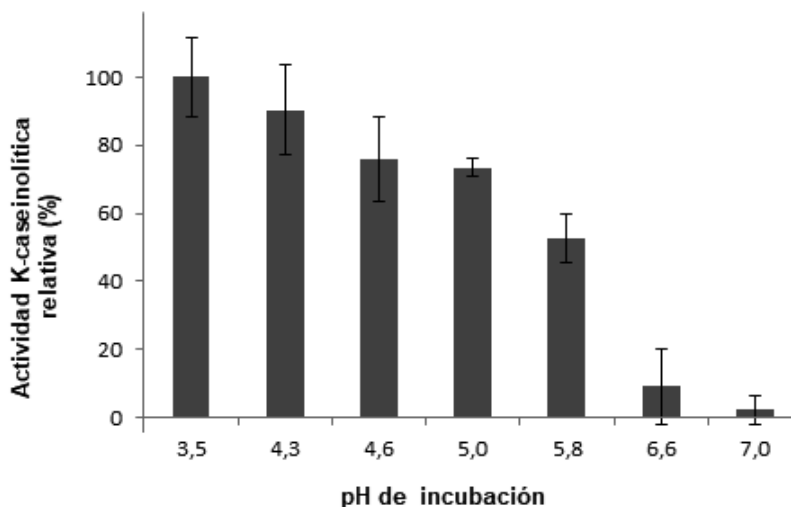


Figura 62. pH de activación de la progaline B. Actividad κ -caseinolítica de la progaline B presente en el medio YPD7 tras haber sido incubada durante 270 minutos a diferentes pHs. La actividad se expresó como un porcentaje de la actividad máxima.

4.5.5 Efecto de la temperatura y el pH sobre la actividad

Determinación de la estabilidad a diferentes temperaturas

El efecto de la temperatura en la estabilidad de la progaline B se muestra en la Figura 63. La enzima fue incubada en un rango de 45 a 90 °C durante 10 ó 30 minutos antes de medir su actividad remanente y compararla con la que presentaba sin ser sometida a ningún tratamiento térmico. La enzima retenía al menos el 90 % de su actividad cuando era incubada entre 45 °C y 55 °C durante 10 ó 30 minutos. Temperaturas superiores a 55 °C daban lugar a una pérdida progresiva de actividad; el tratamiento a 70 °C durante 30 minutos, o a 75 °C durante 10 minutos provocaba la inactivación de la enzima.

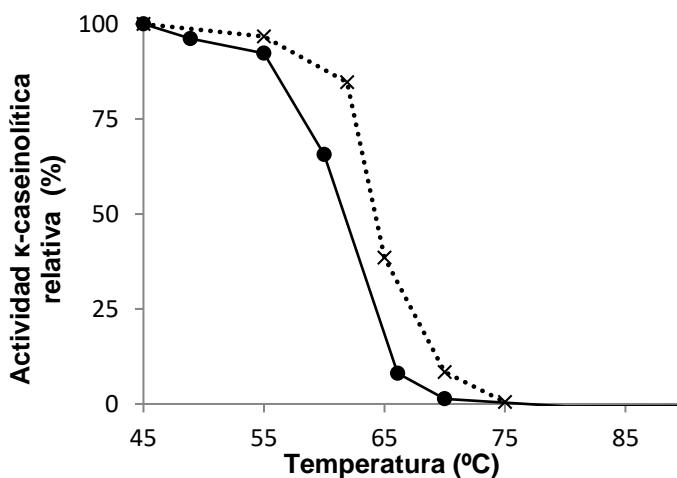


Figura 63. Termoestabilidad de la progeline B. La influencia del tratamiento térmico durante 10 (●●X●●) ó 30 (-●-) minutos en la estabilidad de la progeline B fue determinada en temperaturas en el rango de 45 a 90 °C. La actividad κ-caseinolítica se midió a pH 5.5 e incubación de 1 hora a 37 °C. La actividad de la proteasa sin tratamiento térmico (temperatura ambiente) correspondió al 100 % de actividad, el resto fue expresada como un porcentaje de este valor.

Determinación de la temperatura óptima

La influencia de la temperatura en la actividad proteolítica de la progeline B se muestra en la Figura 64. La actividad óptima de la progeline B fue 50 °C utilizando κ-caseína como sustrato en tampón a pH 5.5. La actividad se incrementaba con la temperatura en el rango que va desde 30 a 50 °C. La actividad relativa era 50 % a 30 °C y 69 % a 35 °C, estas temperaturas son las que normalmente se utilizan en el proceso de elaboración de quesos.

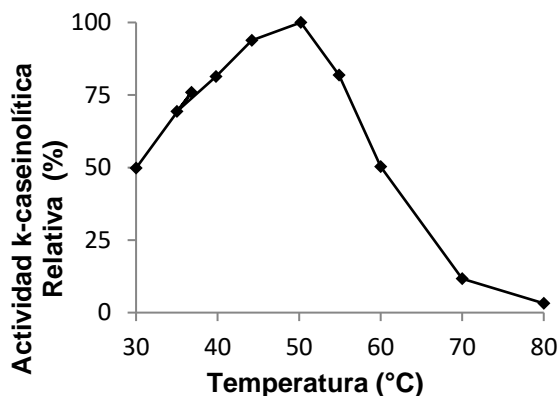


Figura 64. Efecto de la temperatura sobre la actividad de la progaline B. La actividad fue determinada mediante el ensayo FTC- κ -caseína utilizando el tampón acetato sódico 100 mM pH 5.5 y un tiempo de incubación de 1 hora a la temperatura indicada en el gráfico. La actividad se expresó como un porcentaje de la muestra que presentaba la mayor actividad enzimática.

Determinación del pH óptimo

El efecto del pH sobre la actividad de la progaline B se muestra en la Figura 65. La actividad máxima se detectó a pH 5, decreciendo a medida que el pH de la reacción era más alcalino.

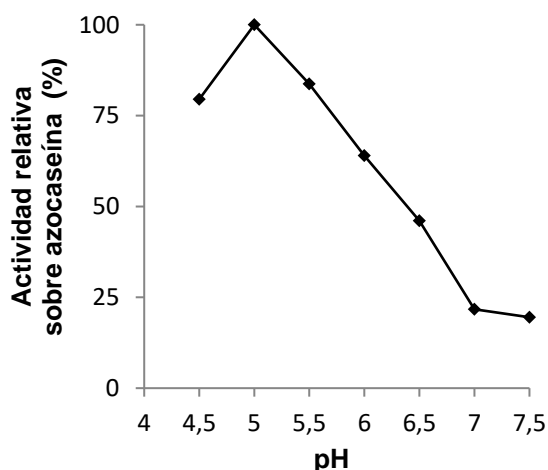


Figura 65. Efecto del pH sobre la actividad de la progaline B. La actividad de la progaline B purificada (9 μ g) fue determinada utilizando 1 % de azocaseína como sustrato en tampones a diferentes pH, la incubación se realizó a 37 °C durante 2 horas. La actividad se expresó como un porcentaje de la muestra que presentaba la mayor actividad enzimática.

4.5.6 Hidrólisis de α -, β - y κ -caseínas

El patrón de hidrólisis de las caseínas es un factor importante para valorar la posible aplicación industrial de un coagulante, debido a que influirá en el rendimiento, textura, sabor y aroma de los quesos elaborados con ella (Fox et al. 2017). Para determinar el grado de hidrólisis efectuado por la progaine B purificada, sobre los diferentes tipos de caseínas bovinas, y compararlo con el efectuado por otros coagulantes de la leche utilizados industrialmente, las enzimas fueron incubadas con α -, β - o κ -caseína independientemente y los péptidos resultantes de la actividad proteolítica visualizados mediante electroforesis en gel de poliacrilamida (Figuras 66 y 67). La fuerza de los diferentes coagulantes fue determinada utilizando leche desnatada suplementada con calcio (Tabla 18) y se calculó el volumen que debería ser añadido a la mezcla de reacción de hidrólisis para que fuese equivalente en todas ellas.

Tabla 18. Fuerza coagulante de la progaine B y preparaciones comerciales.

Coagulante	MCA
Progaine B purificada (0.12 mg/mL)	1:20
Quimosina pura (1 mg/mL)	1:6315
ChyMax® Plus	1:16216
Cuajo estándar de ternera	1:10909
Cuajo estándar de vaca	1:20833
Mucorzyme "L"	1:54794
ChyMax® M	1:57150

La actividad proteolítica de la progaine B sobre la κ -caseína generaba péptidos del mismo tamaño que los obtenidos mediante otros coagulantes comerciales. Así, en la Figura 66.A podemos observar cómo el patrón de bandas obtenido a partir de la digestión de la κ -caseína transcurrida una hora de incubación con la progaine B (calle 3) es el mismo que el generado con otros coagulantes de diferentes orígenes, como son, el cuajo de ternero (calles 4 y 6) y bovino adulto

(calle 7) (origen animal), la quimosina bovina o quimosina de camello producidas en *A. niger* (FPC) (calles 5 y 9 respectivamente) o el coagulante obtenido a partir de *R. miehei* denominado Mucorzyme (origen microbiano) (calle 8); la digestión de la κ -caseína por la progaine B era inhibida cuando la enzima era preincubada con la pepstatina A (calle 2), siendo el patrón de bandas el mismo que el observado para la κ -caseína sin ningún tipo de tratamiento (calle 10). El tamaño del péptido mayoritario obtenido de la hidrólisis de la κ -caseína fue de 13.5 kDa, lo cual es consistente con el tamaño de la para-k-caseína, péptido resultante de la rotura del enlace Phe105-Met106 implicado en la fase primaria de coagulación de la leche (Guinee y O'Brien 2010). La progaine B presentaba una menor especificidad frente a las α - y β -caseínas (Figura 66.B y 66.C, respectivamente, calle3). La hidrólisis de estas caseínas por la progaine B produce péptidos adicionales, que no se observan con los otros coagulantes comerciales analizados, esta hidrólisis es más pronunciada sobre las α -caseínas que sobre las caseínas de tipo β . Dentro de los coagulantes comerciales, el cuajo procedente de bovino adulto y el coagulante procedente de *R. miehei* fueron los que presentaban una mayor actividad proteolítica sobre la α -caseína.

También se estableció la relación entre la actividad coagulante y la actividad proteolítica del sobrenadante concentrado (50X) de la cepa PROGVB-7 utilizando como sustrato la azocaseína. La relación C/P presentaba un valor de 0.06, muy inferior al mostrado por una solución de quimosina pura (10 mg/ml), cuyo valor fue de 10.57.

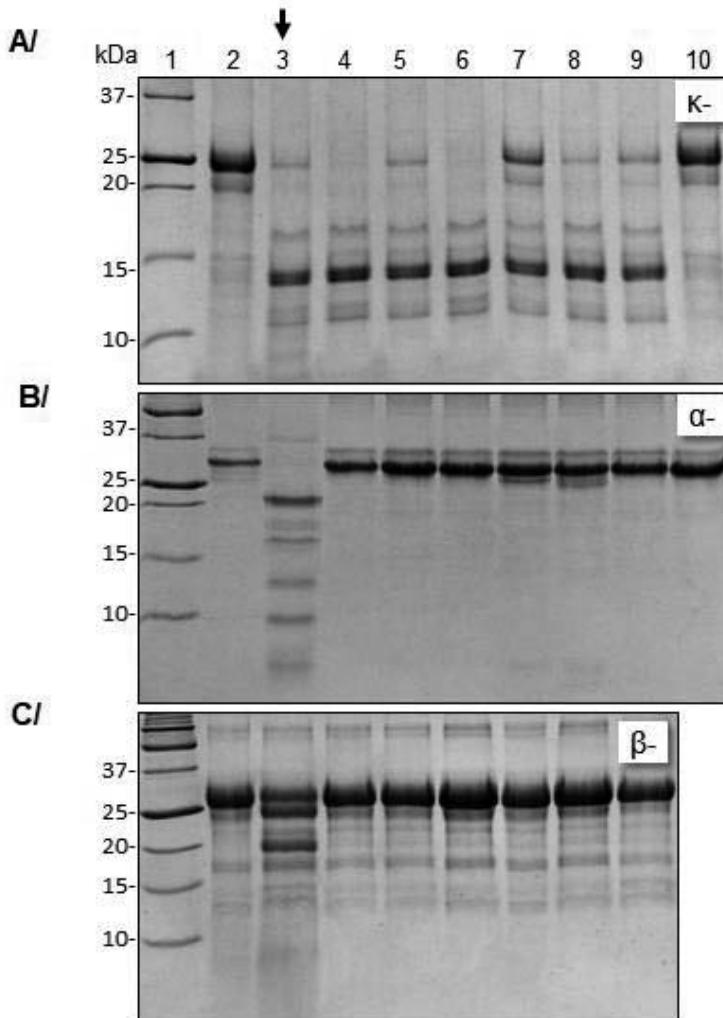


Figura 66. Análisis mediante SDS-PAGE 16.5 % de la hidrólisis de α -, β - y κ -caseínas mediante diferentes coagulantes de la leche (tiempo de incubación 1 hora). (A) κ -caseína, (B) α -caseína y (C) β -caseína fueron incubadas durante 1 hora con progalline B purificada (calle3; indicada con una flecha), quimosina (Sigma; calle 4), quimosina bovina recombinante (ChyMax® Plus; calle 5), cuajo de ternera (calle 6) y vaca (calle 7) (CHR Hansen), coagulante microbiano (Mucorzyme, calle 8), quimosina de camello recombinante (ChyMax® M, calle 9). La calle 1 contiene el marcador de pesos moleculares, indicados en el margen izquierdo. En la calle 2 la progalline B fue preincubada con pepstatina A y en la calle 10 se encuentran la κ -caseína (A) y α -caseína (B) no tratadas.

El patrón de hidrólisis obtenido cuando la digestión se prolongaba durante dos horas era prácticamente el mismo (Figura 67).

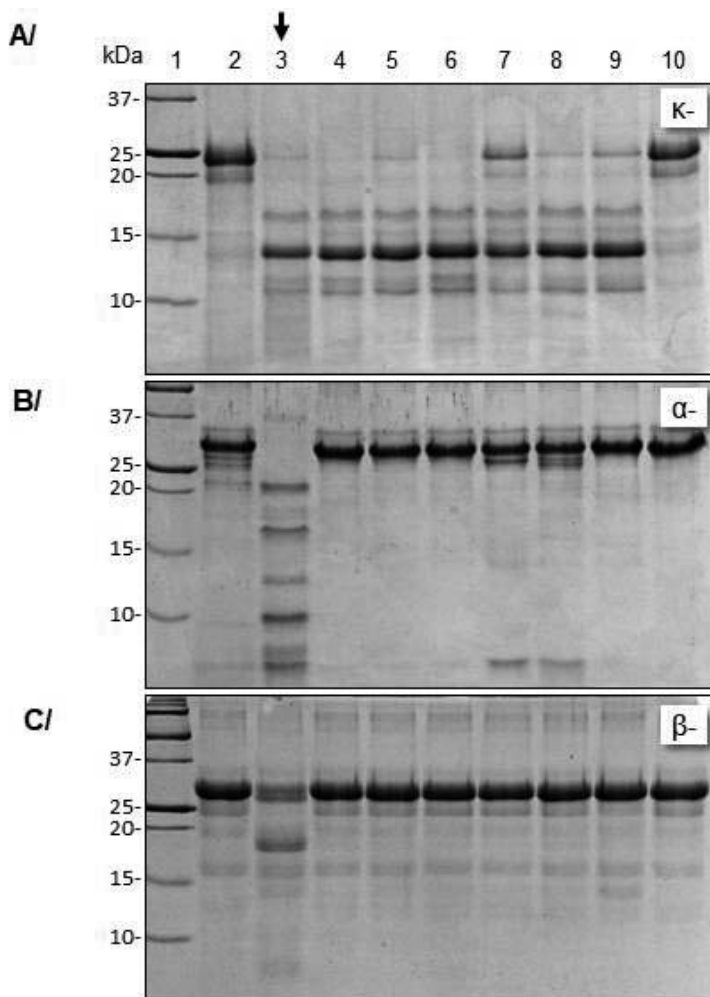


Figura 67. Análisis mediante SDS-PAGE 16,5 % de la hidrólisis de α -, β - y κ -caseínas mediante diferentes coagulantes de la leche (tiempo de incubación 2 horas). (A) κ -caseína, (B) α -caseína y (C) β -caseína (tiempo de incubación 2 horas). Esta figura es homóloga a la Figura 67, salvo que la hidrólisis se prolongó durante 2 horas. En el gel C/, en la calle 10 aparece la β -caseína sin tratar.

4.5.7 Conservación del coagulante

Los aditivos empleados para la conservación de muchos coagulantes de la leche comerciales son benzoato sódico (E-211) y sal.

Se evaluó si el conservante formado por una mezcla de benzoato sódico 0.5 % (E-211) y sal 15 % era adecuado para el almacenamiento del coagulante constituido por el sobrenadante de *P. pastoris* concentrado (50x) conteniendo la progaine B. La preparación se mantuvo libre de microorganismos y sin pérdida de fuerza coagulante durante al menos 6 meses cuando la temperatura de almacenamiento fue de 4 °C.

4.6 Obtención de una cepa superproductora de progaine B

La construcción de una cepa conteniendo múltiples copias del gen de la progaine B se abordó incrementando la concentración de zeocina en las placas de selección de los transformantes (Lin-Cereghino y Lin-Cereghino 2007). En el proceso de transformación de las células competentes de *P. pastoris* se pueden incorporar varias copias del casete de expresión en el genoma celular por integraciones sucesivas del vector. El crecimiento a altas concentraciones de zeocina es indicativo de la incorporación en el genoma de la levadura de varias copias del vector de expresión.

Para incrementar la dosis génica se obtuvieron células competentes de la cepa utilizada hasta el momento (PROGVB-7) y se transformaron con la construcción pGAPZ α -progv2AP, siguiendo el protocolo habitual, salvo que la selección de las cepas se realizó en placas con una concentración de zeocina de 2000 $\mu\text{g/mL}$, muy superior a la indicada para la selección rutinaria (100 $\mu\text{g/mL}$) (Daly y Hearn 2005). La actividad sobre FTC- κ -caseína de los sobrenadantes (20 μL) obtenidos a partir del cultivo de los transformantes en medio YPD durante 3 días no mostró diferencias significativas ($p > 0.05$) con

el sobrenadante de la cepa utilizada para la mejora genética (PROGVB-7) cultivada en las mismas condiciones (230 rpm, 28 °C y DO_{600} inicial= 0.02) (Figura 68). La densidad óptica de los cultivos en el momento del análisis no era exactamente la misma a pesar de utilizar inóculos con la misma densidad celular y realizar el cultivo en las mismas condiciones.

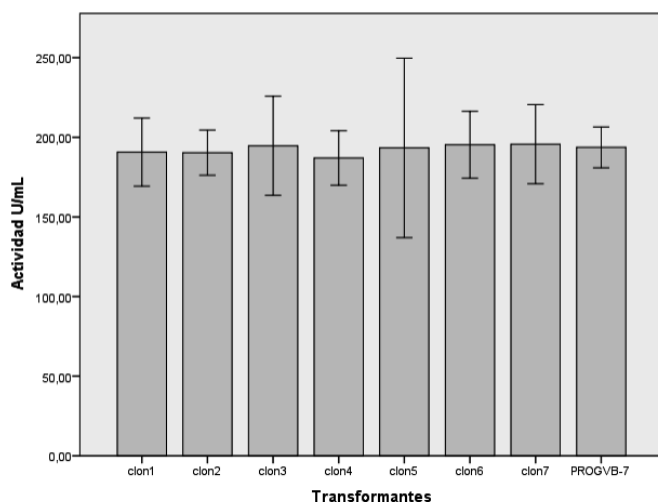


Figura 68. Análisis comparativo de la actividad kappa caseinolítica de los transformantes obtenidos tras un proceso de mejora genética. La cepa PROGVB-7 fue transformada con la construcción pGAPZ α -progv2AP y la selección se realizó utilizando una concentración elevada de zeocina (2000 μ g/mL).

Para determinar si el nivel de expresión podía estar limitado por el uso de codones de la levadura *P. pastoris* se realizó un análisis utilizando el programa Codon Usage Analyser 2.0 (Figura 69). La eficiencia de la traducción podría estar comprometida por la existencia en la secuencia de la progaline B de dos codones, CGG y CGC, que codifican para el aminoácido arginina (posiciones 118 y 299 de la progaline B recombinante), y cuya frecuencia relativa de uso es inferior al 10 %, otros dos codones en las posiciones aminoacídicas 153 y 443 presentaban una frecuencia inferior al 20 %.

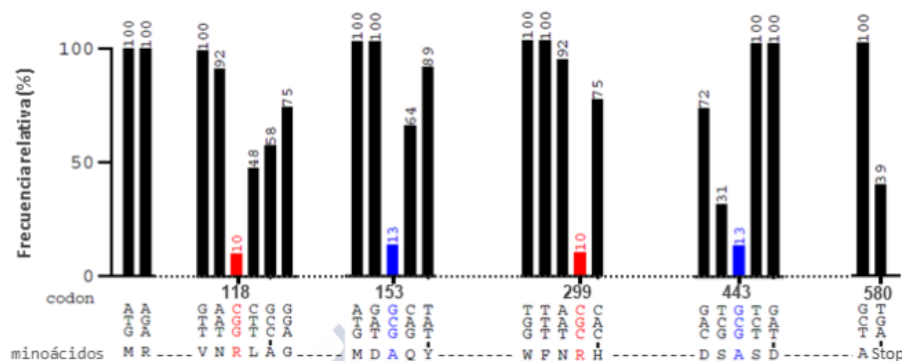


Figura 69. Predicción de la frecuencia de uso de codones raros del gen *gv2AP* cuando se expresa en *P. pastoris*. En el eje horizontal se indica la posición de los codones con una frecuencia de uso baja (inferior al 20 %, columnas en color azul) o muy baja (inferior al 10 %, columnas en color rojo), bajo el eje se indican los codones y aminoácidos que codifican, los números indican la posición de los aminoácidos en la secuencia; la frecuencia de uso de codones (eje vertical) se expresa como porcentaje del codón de frecuencia más alta para cada aminoácido (Sharp y Li 1987).

El índice de adaptación de codones (CAI), obtenido mediante el programa OptimumGene™-Codon Optimization, de la compañía GenScript®, que indica el grado en que el uso de codones de una secuencia se adapta a las frecuencias de uso de codones calculadas previamente para un organismo y utilizado para predecir el nivel de expresión de un gen, fue inferior al 0.74, el óptimo se encuentra entre 0.8-1.0, esto indicaría que la expresión de la progaline B es susceptible de ser mejorada mediante la optimización de los codones al uso del hospedador.

4.7 Evaluación de los quesos elaborados con progaline B

Las características de los quesos producidos con la progaline B fueron evaluadas realizando dos tipos de elaboraciones: un queso

fresco con leche cruda de vaca y un queso curado con leche de oveja pasteurizada suplementada con un cultivo iniciador mesófilo.

Como coagulante se utilizó el sobrenadante concentrado de la cepa recombinante productora de progeline B (cepa PROGVB-7), sin ser sometido a ningún proceso de purificación, ya que previamente comprobamos, en un ensayo realizado con un volumen de 1 mL de leche de vaca suplementada con cloruro cálcico 5 mM, que el sobrenadante concentrado obtenido a partir de la cepa silvestre en las mismas condiciones que la recombinante no tenía capacidad para coagular la leche. Así, mientras el sobrenadante (40 μ L) de la cepa recombinante concentrado (69x) producía la coagulación en 38 minutos (US 1:26), cuando se empleaba el mismo volumen de sobrenadante de la cepa silvestre (70x), la coagulación no se producía, al menos, en las seis primeras horas de incubación a 35 °C.

4.7.1 Queso fresco de vaca

Se elaboraron dos quesos utilizando leche cruda de vaca, uno utilizando como coagulante el sobrenadante concentrado de la cepa PROGVB-7 y otro utilizando quimosina bovina producida por fermentación de *A. niger* (Chymax® Plus, CHR Hansen), procurando que el proceso de elaboración fuese idéntico en ambos casos (Figura 70; este proceso se detalla en el apartado de materiales y métodos 4.7.1).



Figura 70. Imágenes del proceso de elaboración de queso con leche de vaca utilizando como coagulante la progaine B recombinante.

El peso del queso obtenido a partir de 3 L de leche de vaca utilizando como coagulante 100 mL de sobrenadante de la cepa PROGVB-7 concentrado 69X fue de 477.5 gramos, mientras que el peso del queso elaborado con la quimosina producida por fermentación (0.15 mL) fue 462.5 gramos.

Transcurrida una semana se evaluó la influencia del coagulante utilizado sobre las propiedades texturales (definidas en la Tabla 19) y nutricionales de los quesos obtenidos (Figura 71).

El análisis de perfil de textura (TPA), que imita las condiciones a las cuales se somete un material durante el proceso de masticación, se llevó a cabo con un texturómetro TA-XT2i Stable Micro Systems. Se realizaron tres réplicas para cada uno de los quesos y se compararon las medias mediante una prueba t de Student para muestras independientes.



Figura 71. Quesos de leche cruda de vaca elaborados con Chymax® Plus (izq.) y con progaline B recombinante (dcha.).

El análisis de la varianza de los resultados obtenidos para los quesos elaborados con la progaline B y con la quimosina bovina (Tabla 20) indicaba que las propiedades mecánicas analizadas mostraban diferencias estadísticamente significativas ($p < 0.05$) en todos los parámetros analizados a excepción de la elasticidad.

Así, el queso elaborado con la progaline B presentaba una menor dureza, cohesividad, gomosidad y masticabilidad, pero una mayor adhesividad que el obtenido utilizando quimosina bovina. El tipo de coagulante empleado no afectó significativamente ($p > 0.05$) a la composición nutricional de los quesos (Tabla 21), únicamente el contenido en cenizas (residuo inorgánico que queda después de calcinar la materia orgánica) presentaba diferencias significativas, siendo este mayor en el queso elaborado con quimosina bovina.

Tabla 19. Definición de los parámetros de textura analizados (Szczesniak 2002).

Parámetro	Físico	Sensorial
Dureza	Fuerza necesaria para obtener una deformación determinada	Fuerza necesaria para comprimir una sustancia entre los molares (alimento sólido) o entre la lengua y el paladar (alimento semisólido)
Adhesividad	Trabajo necesario para superar las fuerzas de atracción entre la muestra y una superficie	Fuerza necesaria para retirar el alimento que se adhiere a la boca durante su consumo
Elasticidad	Velocidad a la que un material deformado regresa a su condición inicial después de retirar la fuerza deformante	Grado en que un producto regresa a su forma original una vez ha sido comprimido entre los dientes
Cohesividad	Grado de deformación de un material antes de romperse	Grado en que una sustancia es comprimida entre los dientes antes de romperse
Gomosidad	Energía necesaria para desintegrar un semisólido a un estado listo para ser deglutido (combinación de baja dureza y alta cohesividad)	Densidad que persiste durante la masticación; energía requerida para desintegrar un alimento semisólido a un estado listo para tragar.
Masticabilidad	Energía requerida para masticar un alimento hasta que esté listo para ser deglutido (combinación de dureza, cohesividad y elasticidad)	Tiempo requerido para masticar un alimento para reducirlo a una consistencia adecuada para tragarlo

Tabla 20. Análisis de perfil de textura de quesos elaborados con leche de vaca utilizando como coagulante progeline B o quimosina bovina (Chymax™Plus).

Parámetro	Progaline B media \pm DE	Chymax™Plus media \pm DE	Valor <i>p</i>
Dureza (g)	2983.17 \pm 974.75	6583.49 \pm 810.96	0.008
Adhesividad (g s)	-252.87 \pm 86,94	-99.83 \pm 25.89	0.043
Elasticidad	0.49 \pm 0,15	0.61 \pm 0.59	0.258
Cohesividad	0.171 \pm 0,00	0.24 \pm 0.01	0.001
Gomosidad	507.98 \pm 157,42	1574.29 \pm 278.24	0.004
Masticabilidad	261.79 \pm 161,61	960.67 \pm 232.39	0.013

Los valores representan la media de tres medidas \pm la desviación estándar (DE)

El análisis sensorial fue llevado a cabo por diez personas no expertas, realizando una valoración global, seis de ellas mostraron su preferencia por el queso elaborado con el coagulante vegetal.

Tabla 21. Composición y contenido energético de los quesos elaborados con leche de vaca utilizando como coagulante progaine B o quimosina bovina (Chymax™Plus).

Parámetro	Progaine B (media ± DE)	Chymax™Plus (media ± DE)	Valor <i>p</i>
Extracto seco (g)	42.50 ± 0.77	41.47 ± 0.76	0.176
Proteína (g)	15.29 ± 0.11	15.00 ± 0.27	0.158
Hidratos de Carbono (g)	2.98 ± 0.80	1.79 ± 0.34	0.077
Grasa (g)	21.17 ± 0.29	20.83 ± 0.76	0.519
Saturadas (%)	73.64 ± 0.41	74.09 ± 0.21	0.171
Monoinsaturadas (%)	20.64 ± 0.10	20.53 ± 0.08	0.204
Poliinsaturadas (%)	5.61 ± 0.16	5.38 ± 0.24	0.246
Cenizas (g)	3.06 ± 0.18	3.85 ± 0.21	0.008
Calorías (kcal)	263.57 ± 1.05	254.65 ± 6.26	0.072

Composición por cada 100 g de queso

4.7.2 Queso madurado de oveja

Se realizó una segunda masa quesera con la progaine B a partir de leche de oveja pasteurizada suplementada con un cultivo iniciador mesófilo, en este caso no se realizó un control con quimosina a efectos comparativos. El queso se sometió a un proceso de maduración en una quesería artesanal, a una temperatura entre los 12 y 14 °C y una humedad del 95 % durante 50 días, transcurridos los cuales se realizaron los análisis de perfil de textura y nutricionales. En la Figura 72 puede apreciarse el aspecto exterior e interior del queso tras el periodo de maduración.



Figura 72. Queso de oveja elaborado con progaline B transcurridos 50 días de maduración.

Tanto los parámetros texturales analizados (Tabla 22), como la composición nutricional y calórica (Tabla 23), fueron significativamente diferentes a los obtenidos para la elaboración realizada con la progaline B a partir de leche de vaca (Tablas 20 y 21). El queso madurado de oveja, como era de esperar, presentaba una dureza mucho mayor que el queso fresco elaborado con leche de vaca, así como una mayor gomosidad y adhesividad mientras que la cohesividad y elasticidad eran menores ($p < 0.05$), únicamente el contenido en cenizas, hidratos de carbono y la masticabilidad no presentaba diferencias significativas.

Tabla 22. Análisis de perfil de textura de queso elaborado con leche de oveja utilizando como coagulante proguline B.

Parámetro	Proguline B (media \pm DE)
Dureza (g)	17737.46 \pm 4447.98
Adhesividad (g s)	-450.43 \pm 32.23
Elasticidad	0.16 \pm 0.03
Cohesividad	0.12 \pm 0.00
Gomosidad	2176.82 \pm 478.24
Masticabilidad	346.88 \pm 51.97

En cuanto a la composición nutricional, el queso de oveja presentaba una mayor cantidad de extracto seco, proteínas, grasas y un mayor contenido calórico que el queso fresco de vaca elaborado con la proguline B, mientras que el contenido en hidratos de carbono no presentaba diferencias significativas ($p > 0.05$). La evaluación sensorial global realizada por personal no experto fue positiva, apreciando una gran intensidad de sabor, el cual se percibía ligeramente amargo.

Tabla 23. Composición y contenido energético del queso elaborado con leche de oveja utilizando como coagulante proguline B.

Parámetro	Proguline B (media \pm DE)
Extracto seco (g)	76.04 \pm 1.23
Proteína (g)	32.04 \pm 2.61
Hidratos de Carbono (g)	3.90 \pm 0.23
Grasa (g)	35.02 \pm 3.20
Saturadas (%)	68.73 \pm 1.15
Monoinsaturadas (%)	24.30 \pm 1.38
Poliinsaturadas (%)	6.97 \pm 0.62
Cenizas (g)	5.07 \pm 1.72
Calorías (kcal)	458.98 \pm 23.91

Composición por cada 100 g de queso



DISCUSIÓN



5. DISCUSIÓN

Las enzimas activas en todos los coagulantes de la leche comerciales empleados para la fabricación de quesos pertenecen al grupo de las proteasas aspárticas, que pueden ser de origen animal, microbiano, recombinantes o vegetales (Jacob et al. 2011). Extractos de plantas conteniendo diversos tipos de proteasas se han empleado desde la antigüedad para la manufactura de quesos, pero hoy en día su uso está limitado a nivel local (Shah et al. 2014), destacando los extractos procedentes de las flores secas del cardo (Roseiro et al. 2003). El uso de estos coagulantes alternativos favorece la obtención de quesos diferentes a los que habitualmente encontramos en el mercado, pero presentan también una serie de inconvenientes, como son: su elevada actividad proteolítica en relación con la actividad coagulante, el coste que supone la recolección del material vegetal y la estandarización de la calidad.

En este estudio abordamos la búsqueda de proteasas aspárticas de origen vegetal para ser utilizadas como coagulantes de la leche, y su producción en microorganismos, subsanando los inconvenientes asociados al uso de extractos vegetales citados anteriormente. Las especies seleccionadas fueron *Galium verum* y *Ficus carica*, en base a su utilización en el pasado para la producción de quesos (Columela, De Re Rustica; Grieve 1971).

Los extractos elaborados con las hojas y flores de *G. verum* presentan actividad coagulante de la leche, y esta es desencadenada por proteasas de tipo aspártico. La actividad coagulante, al igual que ocurre con la especie vegetal *C. cardunculus*, está presente en extractos elaborados a partir del material vegetal fresco o seco (Chen et al. 2003a). Partiendo de la misma cantidad de material vegetal, el tiempo de coagulación es menor cuando la planta se somete a un

proceso de secado. Las proteasas aspárticas presentes en los extractos frescos y secos podrían ser diferentes, al igual que ocurre con las obtenidas a partir de las flores del cardo, cardosinas y ciprosinas (Heimgartner et al. 1990; Verissimo et al. 1996). La actividad coagulante fue mucho menor de la que esperábamos obtener, a pesar de utilizar un gran volumen de extracto con relación al volumen de leche; el tiempo de coagulación fue superior a dos horas en las condiciones habituales de elaboración del queso. Este resultado parece indicar que, en el pasado, esta especie vegetal se empleó en combinación con coagulantes de otros orígenes.

En el caso de la especie vegetal *F. carica* se verificó que, como ya estaba descrito, el látex presenta actividad coagulante de la leche (Bornaz et al. 2010; Faccia et al. 2012), y esta es debida fundamentalmente a la actividad proteolítica de cisteín-proteasas.

A pesar de la baja actividad coagulante de la leche de los extractos de *G. verum* y de la atribución de la actividad coagulante de *F. carica* a cisteín-proteasas, abordamos la búsqueda de genes codificantes de proteasas aspárticas de ambas especies. Identificamos la secuencia de dos genes de *G. verum* (*gv1AP* y *gv2AP*) y un gen de *F. carica* (*fc1AP*) que codificaban para PAs que denominamos preprogaline A, preprogaline B y preprofigain A respectivamente; también se identificaron dos secuencias parciales de *F. carica* (*fc2AP* y *fc3AP*), todas ellas pertenecían a la familia A1 de la base de datos MEROPS (Rawlings et al. 2018). Estas secuencias son las únicas descritas para proteasas aspárticas de estas especies hasta el momento. Las secuencias se obtuvieron mediante la técnica RACE PCR, utilizando unos cebadores diseñados en base a dos regiones conservadas en la secuencia de un gran número de PAs clasificadas como típicas (Figura 18), por ello los genes identificados codificaban para este tipo de proteasas (Faro y Gal 2005). Al igual que ocurre en otras especies vegetales existen varios genes que codifican para PAs de este tipo en ambas especies (D'hondt et al. 1997; Chen et al. 2002;

Laloi et al. 2002; Sarmiento et al. 2009). El cribado de un mayor número de transformantes y el diseño de cebadores en base a otras regiones conservadas nos permitiría identificar un mayor número de genes. A pesar de que este tipo de proteasas son las mejor caracterizadas dentro de las PAs de origen vegetal, análisis *in silico* indican que es el grupo minoritario (Faro y Gal 2005; Chen et al. 2009; Guo et al. 2013), el diseño de cebadores utilizando secuencias de PAs tipo nucelina o atípicas nos permitiría identificar secuencias de PAs de este tipo en ambas especies.

El análisis del ADN genómico del gen *gvIAP* nos permitió determinar la estructura de la región final del mismo (Figura 22). Los seis exones identificados en el gen *gvIAP* eran interrumpidos por cinco intrones en las mismas localizaciones que presentan otras PAs de plantas de este tipo (Asakura et al. 1995; Cruz de Carvalho et al. 2004; Pimentel et al. 2007; Timotijević et al. 2010), pero a diferencia de estas carecía del intrón situado entre los exones 11 y 12 (o exones 12 y 13 si en la numeración tenemos en cuenta el intrón presente en el extremo 5'UTR, típico de la mayoría de PAs típicas de plantas y que no está presente en las de otros orígenes)(Figura 23). La ausencia de uno de los 12 intrones que normalmente están presentes en los genes vegetales de PAs típicas (13 si tenemos en cuenta el presente en la región 5'UTR) ya ha sido descrita para los genes *AtPasp A2* y *AtPasp A3* de *A. thaliana*, el intrón ausente se encuentra en una localización diferente en cada uno de estos genes, y al igual que el gen *gvIAP*, los otros intrones se encuentran en posiciones idénticas al resto de las PAs de plantas descritas (Chen et al. 2002). El tamaño de los intrones variaba en todas las especies utilizadas para realizar el estudio comparativo. El hecho de que la localización intrón-exón sea idéntica en monocotiledóneas y dicotiledóneas (Figura 23) parece indicar que la “inserción” de los intrones ocurrió antes de la divergencia de ambas clases de plantas, mientras que la pérdida o ganancia de nucleótidos correspondientes a cada intrón, o la ausencia de alguno de los

intrones, habría tenido lugar después de la divergencia, explicándose así la longitud variable de los intrones de PAs típicas de genes pertenecientes a la misma especie (Pimentel et al. 2007), así como el hecho de que en alguna especie como *A. thaliana* alguno de sus genes carezca de un intrón mientras que otros lo mantienen (Chen et al. 2002). Los genes que codifican para PAs de vertebrados también presentan una estructura de exones-intrones conservada, consistente en 9 exones y 8 intrones localizados en posiciones similares, mientras que algunos de los genes de PAs de levaduras no presentan intrones y los pertenecientes a otros grupos no tienen una organización estructural conservada (Morales et al. 2004).

Las secuencias de aminoácidos deducidas a partir de los ADNc secuenciados mostraron una elevada homología con otras PAs típicas de varias especies vegetales como la fitepsina de cebada (Runeberg-Roos et al. 1991), la StAP de la patata (Guevara et al. 2005), *SoyAIP* de la soja (Terauchi et al. 2004), la AtAP1 de *A. thaliana* (Chen et al. 2002; Mazorra-Manzano y Yada 2008) o la cardosina A (Faro et al. 1999), cardosina B (Vieira et al. 2001) y ciprosina B (Sampaio et al. 2008) pertenecientes a la especie *C. cardunculus* (Figura 25). La secuencia de aminoácidos correspondientes al prosegmento y al PSI presenta un menor grado de homología entre las diferentes especies que los extremos N- y C-terminal correspondientes a la enzima madura. En el caso del prosegmento la causa de la baja homología podría ser debida al hecho de que los contactos establecidos entre el prosegmento y la región de la enzima madura son débiles y únicamente se conservarían aquellos residuos que desempeñan un papel importante en el mantenimiento de la estructura. La región del PSI que presenta menor homología es la que corresponde a la secuencia situada entre los aminoácidos 40 al 65 (situada entre los dominios Sap_B1 y Sap_B2) (Figura 27), región que no está presente en las saposinas animales (Bliven y Prlić 2012); el PSI podría derivar de la región central de un gen ancestral similar al precursor de las

saposinas (prosaposina), en concreto del extremo C-terminal de una saposina, la región linker y la región N-terminal de la siguiente saposina (Figura 27) (Ponting y Russell 1995; Bliven y Prlić 2012), la región con menor homología correspondería a esa región linker que no estaría presente en las saposinas maduras.

La homología con la quimosina bovina, utilizada para la fabricación del queso, es también elevada si se omite la secuencia correspondiente al PSI al realizar el alineamiento (Figura 25). El PSI únicamente ha sido identificada en plantas, no posee homología con ninguna secuencia de PAs de animales o microorganismos, sin embargo, posee una gran similitud con un grupo de proteínas denominadas SAPLIP (Bruhn 2005).

El análisis *in silico* de las secuencias obtenidas mostró que todas ellas presentaban los motivos estructurales propios de las PAs típicas de plantas (Figura 26), un péptido señal, seguido de un prosegmento de aproximadamente 40 aminoácidos y un fragmento específicos de plantas situado entre los extremos N- y C-terminal, con una longitud de ~100 aminoácidos (Simões y Faro 2004) (Dunn 2002). Las tres PAs tenían en el extremo N-terminal residuos equivalentes a Lys11 y Tyr13 de la fitepsina, conservados en la mayoría de las plantas, exceptuando las cardosinas y una PA de *C. humilis* (Pimentel et al. 2007); estos residuos estarían implicados en el mecanismo de inactivación de las PAs vegetales de manera análoga a lo que ocurre con los residuos Lys36p y Tyr37 del prosegmento de las proteasas gástricas (Kervinen et al. 1999; Simões y Faro 2004) y según sugieren Koelsch y colaboradores (1994) nos indicaría que las tres PAs identificadas se activarían a pH ácido, lo cual ha sido verificado para la preprogaline B.

La preprogaline B y la secuencia parcial que codifica para la preprofigain C presentan el motivo KGE, situado en posición homóloga al motivo RGD que caracteriza a la cardosina A (Faro et al. 1999), en el dominio N-terminal, entre las secuencias que forman

parte del sitio activo (DTG y DSG). Al igual que la mayoría de las PAs vegetales todas las secuencias identificadas en este estudio presentaban uno o dos sitios potenciales de glicosilación (Costa et al. 1997) y seis residuos de cisteínas conservados en el PSI implicados en la formación de tres puentes disulfuro. El sitio de glicosilación localizado en el PSI característico de la mayoría de las PAs típicas de plantas como la fitepsina de cebada (Kervinen et al. 1999), AcAP1 de la piña (Raimbault et al. 2013), cardosina B, procirsin, ciprosina B, ciprosina A, cenprosina (Vairo-Cavalli et al. 2013) y AtAP A1 (Mazorra-Manzano y Yada 2008), únicamente está presente en la preprogaline B y en la secuencia parcial de la preprofigain B.

Se obtuvieron las estructuras tridimensionales de las PAs mediante modelado por homología (Nielsen et al. 2010) utilizando una de las pocas PAs vegetales cuya estructura cristalográfica ha sido determinada, la fitepsina de *H. vulgare* (Kervinen et al. 1999) sin procesar, conteniendo la región correspondiente al prosegmento y al PSI [en este documento únicamente se muestra la correspondiente a la preprogaline B, por ser las tres estructuras similares y ser esta enzima la utilizada para la elaboración de las masas queseras]. Al igual que otras enzimas de la familia A1 la estructura de estas proteasas es bilobulada, el sitio activo se encuentra entre los dos lóbulos, conteniendo cada uno de ellos uno de los residuos de ácido aspártico, pero a diferencias de las PAs de otros orígenes, en las cuales ambos residuos Asp forman parte de la secuencia Asp-Thr-Gly (Davies 1990) en el caso de las PAs identificadas uno de los residuos se encuentra formando parte de la secuencia Asp-Ser-Gly (Simões y Faro 2004) (Kervinen y Wlodawer 2013), esta característica no es exclusiva de las PAs de plantas, la comparten con algunas PAs de hongos y protozoos (MacKay et al. 1988; Hirata et al. 1988; Lee y Kolattukudy 1995; Olsen et al. 1999; Lin et al. 2000; Li et al. 2010). La implicación fisiológica de esta variación no ha sido determinada.

El análisis filogenético elaborado para establecer la relación evolutiva entre las secuencias obtenidas en este estudio y otras PAs vegetales típicas, mostraba que la preprogaline A se encontraba en un grupo diferente a la preprogaline B y más próxima a la preprofigain B, a pesar de pertenecer a especies diferentes. La preprogaline B presentaba una relación más estrecha con el grupo que las PAs de *C. cardunculus*, especie utilizada en la elaboración de quesos (Roseiro et al. 2003) que con el resto de las PAs analizadas (Figura 31).

Para estudiar la posible aplicación en la elaboración de quesos de las proteasas codificadas por los genes identificados se abordó su expresión. El hospedador seleccionado como primera opción fue la bacteria *E. coli*, la elección de este microorganismo fue motivada por su elevada tasa de crecimiento, alta capacidad de producción, reducido coste de los medios de cultivo y la existencia de un gran número de herramientas moleculares que facilitan su manipulación genética (Rosano y Ceccarelli 2014). Sin embargo, una gran parte de las proteínas eucariotas producidas utilizando este microorganismo se producen incorrectamente plegadas, formando agregados insolubles conocidos como cuerpos de inclusión, debido a que no realiza modificaciones postraduccionales, como glicosilaciones, y el ambiente reductor de su citoplasma dificulta la formación de los puentes disulfuro (Rinas et al. 2017). Pese a ello, este hospedador se ha utilizado durante años para la producción industrial de la quimosina bovina (Pfizer 1988) y ha sido objeto de varias investigaciones para la mejora del sistema de expresión de la misma, así como de quimosinas procedentes de otras especies animales (Kumar et al. 2010). La expresión del gen *gvIAP*, sin la secuencia correspondiente al péptido señal y fusionado a una cola de histidinas, dio lugar a la obtención de la progaline A (~58 kDa) en grandes cantidades, pero como era de esperar, únicamente se detectaba en la fracción insoluble citoplasmática, indicándonos que se producía en forma de cuerpos de inclusión. La mayoría de las PAs, tanto de origen animal (Kumar et al.

2010), como vegetal [(arroz, Asakura et al. 2000), (*C. cardunculus*, Castanheira et al. 2005; White et al. 1999), (batata, Chen et al. 2015), (*Arabidopsis*, Simões et al. 2007)(*C. vulgare*, Lufrano et al. 2012)], expresadas utilizando este hospedador se producen, al igual que la progaine A, en forma de estos agregados insolubles. La estrategia basada en la disminución de la temperatura de cultivo durante la inducción para aumentar la cantidad de la enzima correctamente plegada en relación a la que se encontraba formando parte de estos agregados no tuvo éxito (Sørensen y Mortensen 2005). Para la obtención de la enzima funcional es necesaria la desnaturalización y plegado de estos agregados, para lo cual se han desarrollado varias estrategias (Vallejo y Rinas 2004; Yamaguchi y Miyazaki 2014; Singh et al. 2015). La purificación de la progaine A mediante cromatografía de afinidad en condiciones desnaturalizantes y su posterior plegado, utilizando dos métodos basados en la dilución rápida en una solución Tris HCl 20 mM y ajuste lento a pH 8, no nos permitió obtener la proteasa en su forma activa. Sin embargo estos métodos han sido utilizados con éxito para la obtención de otras proteasas aspárticas funcionales de diferentes orígenes, como la cardosina A (Castanheira et al. 2005), el pepsinógeno porcino, una PA de *Rhizopus chinensis*, la cathepsina D, memapsin 1 y 2, napsin, o la proteasa HIV entre otras (Lin 2006). La temperatura a la que se realiza el ajuste del pH después de la dilución tiene influencia en el proceso de renaturalización de la progaine A, así, cuando se realiza a 4 °C conserva el tamaño que presenta en los agregados insolubles, mientras que cuando se efectúa a temperatura ambiente, la proteasa sufre un procesamiento, la banda de ~58 kDa, observada tras la purificación y desnaturalización de los cuerpos de inclusión, pasa a tener un tamaño de ~34 kDa (Figura 36), muy similar al que presenta la cadena pesada de la cardosina A en su forma activa (Castanheira et al. 2005). Esto parecía indicar que el método aplicado a la cardosina A era reproducible para la progaine A, sin embargo, tras su incubación a pH

ácido para su activación, la enzima no presentó actividad proteolítica. El uso de la cepa Rosetta-gami (DE3) para la expresión de la progaline A y profigain A tampoco nos permitió la producción de estas en forma soluble, a pesar de que esta cepa favorece la formación de puentes disulfuro y facilita la expresión de proteínas eucariotas; la estrategia basada en la incubación de la cepa a una temperatura inferior a la óptima durante la inducción tampoco tuvo éxito. La presencia de seis puentes disulfuro en la estructura de ambas proteasas probablemente imposibilite su obtención en forma soluble utilizando un hospedador procariota. Debido al impacto negativo que tiene la producción de las proteasas recombinantes en forma de cuerpos de inclusión para el proceso a nivel industrial abordamos la clonación en un hospedador eucariota. Después de evaluar varios microorganismos, entre los que se encontraban la levadura *K. lactis* y el hongo *A.niger*, utilizados desde hace muchos años para la producción de la FPC a escala industrial (Harboe et al. 2010), nos decantamos por la levadura metilotrófica *P. pastoris*, utilizada con éxito para la expresión de otras proteasas aspárticas con actividad coagulante de la leche (White et al. 1999; Jiang et al. 2012; Gama Salgado et al. 2013; Yegin y Fernandez-Lahore 2013; Nosedá et al. 2013; Wang et al. 2015; Luo et al. 2016; Espinoza-Molina et al. 2016). Entre las ventajas que presenta figuran: un metabolismo fundamentalmente respiratorio, lo cual permite que crezca a elevadas densidades celulares, debido a que no genera cantidades significativas de productos tóxicos; capacidad para crecer en medios simples y de bajo coste; secreción eficiente de proteínas heterólogas; bajo nivel de proteínas endógenas secretadas, lo cual facilita la purificación de la proteína recombinante; realiza modificaciones postraduccionales; fácil escalado a biorreactores; manipulación genética sencilla (Ahmad et al. 2014; Cregg et al. 2009); generación de transformantes estables; ha sido reconocido por la FDA (Food and Drug Administration) como organismo GRAS (generally recognized as safe) (Spohner et al. 2015). Nuestro equipo de

investigación ha utilizado con éxito este sistema de expresión para generar varios productos de interés industrial, entre los que figura, la quimosina de búfalo (Vallejo et al. 2012), una proteasa de *Yarrowia lipolytica* (Poza et al. 2007) y varios carotenoides (Araya-Garay et al. 2012).

Todas las construcciones para la transformación de *Pichia* se realizaron sustituyendo el péptido señal nativo de las proteasas de *G. verum* y *F. carica*, por el factor- α de *S. cerevisiae* presente en los vectores utilizados para la expresión, pPICZ α y pGAPZ α (Cregg 2007; Daly y Hearn 2005). La expresión utilizando el promotor inducible por metanol del gen AOX1 de las proteasas preprogaline A y preprofigain A no tuvo éxito, a pesar de que los genes que las codifican se estaban transcribiendo, el patrón de bandas obtenido mediante SDS-PAGE era idéntico en la cepa silvestre y en los transformantes, y el sobrenadante no presentaba actividad enzimática con relación a la cepa silvestre, indicándonos que las proteasas no se estaban expresando (Figura 42). El resultado fue el mismo cuando se realizó la transformación utilizando el vector pGAPZ α para la expresión de estas proteasas bajo el promotor constitutivo del gen gliceraldheído-3-fosfato deshidrogenasa. La preprogaline B fue la única proteasa que conseguimos expresar de manera constitutiva en *P. pastoris* utilizando este último vector. El motivo por el cual las transformaciones realizadas con las construcciones pPICZ α -progv1AP, pPICZ α -profc1AP, pPICZ α -fc1AP, pGAPZ α -progv1AP, pGAPZ α -profc1AP pPGAPZ α -fc1AP no dieron lugar a la expresión de las proteasas lo desconocemos, una de las diferencias fundamentales en el diseño de todas ellas en relación a la construcción pGAPZ α -progv2AP, que dio lugar a la obtención de la progaline B, es que la fusión del factor- α de *S. cerevisiae* se realizó a la secuencia de la progaline B conteniendo la secuencia que codifica para los siete últimos aminoácidos correspondientes al péptido señal nativo de la enzima, mientras que en el resto de las construcciones la fusión

traduccional se realizó entre el factor- α y los nucleótidos que codifican para el primer aminoácido del prosegmento (en el caso de pPICZ α -progv1AP, pPICZ α -profc1AP, pGAPZ α -progv1AP y pGAPZ α -profc1AP) o bien a la secuencia que codifica para el primer aminoácido de la enzima madura (en el caso de pPICZ α -fc1AP y pPGAPZ α -fc1AP) (Figuras 39 y 44). Los primeros aminoácidos presentes en el extremo N-terminal de la enzima secretada podrían jugar un papel fundamental en el correcto plegamiento y en la estabilidad del enzima (Yegin y Dekker 2013). El efecto de la situación de los sitios de N-glicosilación (Figura 27) podría estar implicado también en el éxito de la secreción de las proteasas recombinantes al medio de cultivo, así como en la actividad de las mismas (Jenkins y Curling 1994; Harboe 1998; van den Brink et al. 2006), la preprogaline A y preprofigain A presentan un sitio de glicosilación muy próximo al sitio activo que podría afectar a esta.

La cepa de *P. pastoris* transformada con el gen *gv2AP* secretaba la progaline B al medio de cultivo (YPD), presentando actividad coagulante de la leche sin necesidad de un proceso de activación previo. Se han descrito otras PAs recombinantes secretadas por *P. pastoris* al medio YPD en su forma activa (Vallejo et al. 2008; Gama Salgado et al. 2013; Yegin y Fernandez-Lahore 2013). Esto supone una ventaja frente a otros sistemas de expresión utilizados para la producción de quimosina, como el basado en *K. lactis*, en los que la enzima es secretada al medio con el prosegmento, requiriendo un proceso de activación (Harboe et al. 2010). La secreción de la proteasa parece indicar que el PSI no juega un papel importante en el direccionamiento hacia las vacuolas, al igual que ocurre con las proteasas recombinantes cipsosina y rAtAPA1 (White et al. 1999; Mazorra-Manzano y Yada 2008).

La enzima recombinante activa está constituida por dos cadenas polipeptídicas, una cadena pesada de 30.7 kDa y una cadena ligera de 13.5 kDa, unidas mediante puentes disulfuro. En el proceso

postraduccional se produce la N-glicosilación del sitio potencial de glicosilación predicho, localizado en el PSI; al igual que ocurre con otras PAs expresadas en *Pichia*, la enzima se secreta al medio en su forma glicosilada y no glicosilada (Gama Salgado et al. 2013). Los oligosacáridos unidos al sitio de N-glicosilación de la progaine B son de alto contenido en manosas, pudiendo ser hidrolizados por la endoglicosidasa H, enzima que, a diferencia de la PNGaseF, no elimina oligosacáridos complejos. La disminución del peso molecular de la progaine B después del tratamiento con la endoglicosidasa H es reducido (~5 kDa), esto también se observa en otras PAs de origen animal y microbiano producidas en *P. pastoris* (Wang et al. 2015; Gama Salgado et al. 2013; Vallejo et al. 2008), estas glicoproteínas son sintetizadas con oligosacáridos que constan de 8 a 14 residuos de manosa por cadena (Grinna y Tschopp 1989), mucho más cortos que los introducidos por *S. cerevisiae* (50-150 residuos de manosa), lo cual hace que la glicosilación sea semejante en tamaño a los oligosacáridos de los organismos superiores (Cereghino y Cregg 2000; Bretthauer y Castellino 1999). La cadena ligera conserva en su estructura parte del PSI, aunque desconocemos el lugar exacto donde se produce el procesamiento proteolítico de este dominio, sabemos que el fragmento que contiene los aminoácidos correspondientes al sitio de N-glicosilación permanecen en la estructura de la enzima activa (Figura 56).

El proceso por el cual la progaine B es obtenida en su forma activa desde el medio de cultivo (YPD) se determinó mediante el crecimiento de la cepa recombinante en medio tamponado a pH ácido y pH neutro. La proteína recombinante es secretada al medio de cultivo tamponado a pH 7 como una única cadena polipeptídica de 56 kDa, que corresponde al precursor inactivo glicosilado (conteniendo el prosegmento y el PSI), mientras en el medio tamponado a pH 4 la progaine B se encuentra en su forma activa, presentando actividad κ -caseinolítica (Figura 61). El ensayo de activación *in vitro* a diferentes

valores de pH nos permitió determinar que el precursor de la progline B (56 kDa) da lugar a una pequeña cantidad de enzima activa cuando se incuba a ~ pH 6 durante 270 minutos a 28 °C, mientras a pH 3.5 la activación es completa transcurridos 150 min (Figura 61). Esto ocurre también con la proquimosina, la cual necesita permanecer entre 2 y 3 días a pH 5 y temperatura ambiente para su activación, mientras que a pH 2 únicamente necesita 5-10 min (Chitpinitol y Crabbe 1998). En el proceso de activación se detecta un péptido de 37 kDa correspondiente a la cadena pesada de la progline B con el prosegmento. El tratamiento del precursor de la progline B con pepstatina A, un inhibidor específico de proteasas aspárticas impide el procesamiento, indicando que la escisión del prosegmento es un proceso autocatalítico, ya que descartamos que este se produzca por PAs secretadas por *Pichia*, debido a la ausencia de actividad proteolítica del sobrenadante de la cepa silvestre (Figura 59). Este tipo de procesamiento se ha observado en otras PAs vegetales como la cenprosina (Domingos et al. 2000), cardosina A (Castanheira et al. 2005; Ramalho-Santos et al. 1998), fitepsina (Glathe et al. 1998) y rAtAP A1 (Mazorra-Manzano y Yada 2008).

Concluimos así que la progline B es sintetizada y secretada al medio como un precursor inactivo y que la activación se produce debido a la progresiva acidificación del medio durante los primeros días de cultivo como consecuencia de la actividad metabólica de *Pichia* (Figura 58).

La expresión de otras PAs vegetales recombinantes como la ciprosina B (White et al. 1999) y AtAP A1 (Mazorra-Manzano y Yada 2008) en *P. pastoris* también conducen a la obtención de una proporción de las mismas en su forma activa (10-50 % en el caso de la AtAP A1) en el medio de cultivo, a pesar de estar tamponado a pH 6. En ambos casos los autores especulan que la activación de estas proteasas podría producirse a pH próximo al neutro por la acción de proteasas endógenas secretadas por *Pichia* al medio de cultivo, a pesar

de que en uno de los estudios se empleó la cepa de *Pichia* SMD1168H deficiente en proteasas, una explicación alternativa sería la activación mediada por pH, al igual que ocurre con la progaline B.

Los resultados obtenidos indican que la activación de la progaline B recombinante implica tres pasos: i) la proteasa sufre el procesamiento del factor- α de *S. cerevisiae* que la conduce a través de la ruta de secreción, dando lugar a las formas glicosilada y no glicosilada de la misma, con pesos moleculares de 56 y 51 kDa respectivamente ii) el siguiente paso implica la eliminación parcial del PSI, que ocurre en el medio de cultivo y da lugar a dos cadenas polipeptídicas, una de 37 kDa y otra de 13.5 kDa iii) finalmente el prosegmento sufre un procesamiento autoproteolítico dando lugar a la enzima activa, constituida por dos cadenas polipeptídicas de 30.7 kDa (HC) y 13.5 kDa (LC) unidas mediante puentes disulfuro, la LC es la que se encuentra glicosilada. Este procesamiento muestra diferencias con el descrito para algunas PAs vegetales nativas y recombinantes: i) La forma activa de la progalina B retiene en su estructura parte de los aminoácidos del extremo N- y C- terminal del PSI, a diferencia de lo que ocurre con otras PAs de plantas típicas, como la procarnosina A nativa procesada *in vitro* (Ramalho-Santos et al. 1998), la fitepsina producida en células de insecto (Glathe et al. 1998) o una PA del girasol (Park et al. 2001). El peso molecular de la cadena pesada indica que parte del extremo N-terminal del PSI está presente en la estructura de la enzima activa y la cadena ligera contiene al menos la región del C-terminal en la que se encuentran los residuos Asn-Gln-Thr que corresponden al sitio de glicosilación. ii) La progaline B presenta también diferencias en cuanto al proceso de activación y su orden. Así, la procirsina producida en *E. coli* presenta actividad en su forma monomérica, sin necesidad del procesamiento del prosegmento ni el PSI (Lufrano et al. 2012) y en el caso de la procarnosina A la escisión del prosegmento es anterior al procesamiento del PSI (Castanheira et al. 2005). iii) El procesamiento de otras PAs como la

fitepsina recombinante y la procardosina A, nativa o recombinante, ocurre a través de formas intermedias antes de dar lugar a la enzima activa (Castanheira et al. 2005) (Glathe et al. 1998).

Los resultados obtenidos indican que a pesar de que los precursores de PAs vegetales típicas presentan una estructura primaria que consta de los mismos dominios, los mecanismos de activación presentan numerosas variantes.

El gráfico que relaciona el crecimiento celular en YPD y la actividad enzimática muestra dos perfiles distintos (Figura 58), sugiriendo que existe un límite en la producción de la enzima activa, que estaría asociado a la alcalinización que sufre el medio de cultivo transcurridos tres días, lo cual provoca que el precursor de la progaine B secretado no se active y se acumule en el medio de cultivo en su forma inactiva.

Existen varios factores que influyen en la actividad y estabilidad de los coagulantes de la leche que son importantes a la hora de evaluar su potencial aplicación, entre ellos se encuentran la temperatura y pH óptimos y la termoestabilidad. La actividad enzimática de la progaine B es máxima a una temperatura de 50 °C y a pH 5 (utilizando FTC- κ -caseína y azocaseína como sustratos). Este valor de temperatura óptima es superior al que presentan las quimosinas recombinantes del camello, búfalo, cabra y vaca (rango 37-45 °C) utilizando el mismo sustrato para su determinación (Vallejo et al. 2012) y el valor óptimo de pH se encuentra dentro del rango observado para la quimosina bovina (pH 4.5-5) utilizando los diferentes tipos de caseínas para su determinación (Kumar et al. 2010). Los valores de temperatura y pH óptimos presuntamente presenta la progaine B también son superiores a los observados para PAs de *C. cardunculus*, especie vegetal utilizada en la elaboración de quesos: la cardosina A nativa y recombinante, cuya temperatura óptima se encuentra en el rango de 40-45 °C y el pH óptimo entre 4-4.5 (Castanheira et al. 2005); VRen (cardosina B modificada

recombinante; Almeida et al. 2014) con 40 °C y pH 4.5, y la ciprosina B recombinante con 42 °C y un pH óptimo entre 4.5-5.1 (Sampaio et al. 2008). Estos valores dependen de los sustratos y condiciones del ensayo, por lo que únicamente nos permiten concluir que la temperatura y pH óptimos de la progaine B se encuentra dentro del rango presentado por otros coagulantes de la leche utilizados en la elaboración de queso. Otro factor a tener en cuenta a la hora de evaluar una enzima destinada a la elaboración de quesos es que pueda ser inactivada mediante tratamiento térmico a una temperatura que no afecte significativamente a las proteínas del suero lácteo, para que estas puedan ser utilizadas como subproducto del queso (Harboe et al. 2010). La progaine B cumple este requerimiento, siendo inestable a temperaturas superiores a 55 °C. Sin embargo, no es completamente inactivada mediante el proceso de pasteurización, requiriendo al menos un tratamiento de 10 minutos a 75 °C.

La rotura específica de la proteína láctea κ -caseína, que desencadena la fase primaria de coagulación de la leche, es otra característica deseable en un coagulante de la leche. La progaine B recombinante muestra un patrón de hidrólisis sobre la κ -caseína bovina prácticamente idéntico al que presenta la quimosina bovina (Reid et al. 1997), aunque no podemos asegurar que el enlace hidrolizado sea el mismo sobre el que actúa esta, Phe105-Met106 (Dalglish 1993). La mayor diferencia entre la progaine B y otros coagulantes comerciales se encuentra en el patrón de hidrólisis de las α - y β - caseínas, sobre las que la progaine B ejerce una acción proteolítica mucho más extensa que el resto, siendo la α - caseína más susceptible a la hidrólisis que la β -caseína, al igual que ocurre con las cynarasas y la cardosina B (Sidrach et al. 2005; Almeida et al. 2014). Estos resultados deben ser verificados analizando quesos elaborados con la progaine B, ya que la accesibilidad de los enlaces que son hidrolizados puede variar en las condiciones y entorno que existe en el

queso en relación a los que existen cuando las caseínas se encuentran en solución (Upadhyay et al. 2004; Exterkate et al. 1997).

Una elevada proteólisis de las caseínas de la leche se asocia al desarrollo de sabores amargos (Habibi-Najafi y Lee 1996 ; Lemieux y Simard 1992) y defectos en la textura en el queso, así como a un bajo rendimiento quesero (Harboe et al. 2010). A pesar de ello, las flores secas procedentes de *Cynara* spp., que presentan una elevada acción proteolítica sobre las α - y β - caseínas, son utilizadas para producir quesos muy apreciados por los consumidores, y es justamente la extensiva hidrólisis de estas caseínas, la responsable de que los quesos elaborados con estos coagulantes tengan unas características reológicas y sensoriales muy diferentes a los que presentan los quesos elaborados con los coagulantes actualmente presentes en el mercado (Sousa y Malcata 1997; Roseiro et al. 2003; Chen et al. 2003a; García et al. 2012). Esta característica está siendo investigada para ser utilizada en la maduración acelerada de los quesos, reduciendo así los costes de producción de ciertos tipos de quesos (Prados et al. 2007; Galan et al. 2012).

Los niveles de expresión de la progaline B en *P. pastoris* fueron muy bajos, siendo necesario concentrar ~50 veces el sobrenadante de la cepa productora para que el volumen añadido a la leche produjese la coagulación en un tiempo inferior a una hora. Por ello se abordó la obtención de una cepa hiperproductora retrotransformando la cepa inicial y realizando la selección utilizando cantidades crecientes del factor de selección, para incrementar así la dosis génica (Damasceno et al. 2012). Ninguna de las cepas obtenidas presentaba una diferencia significativa en la secreción de la progaline B con respecto a la cepa original. Esta estrategia para optimizar la producción ya había sido utilizada por el grupo de investigación para la mejora de una cepa de *Pichia* productora de quimosina de búfalo sin éxito (Vallejo et al. 2008). Aunque existen numerosos casos en los que existe una relación directa entre el número de copias del gen y la

secreción de la proteína recombinante (Damasceno et al. 2012; Zhang et al. 2009), se han encontrado casos en los que una mayor dosis génica no implica un aumento en la producción, o incluso puede ir en detrimento de esta debido a un efecto negativo sobre el crecimiento de *Pichia* o sobre la ruta de secreción (Macauley-Patrick et al. 2005; Inan et al. 2006; Zhu et al. 2009; Liu et al. 2014). El bajo nivel de producción podría ser atribuido al sesgo de codones existentes entre *Pichia* y *G. verum* (Yang y Zhang 2018), el gen de la progaine B contiene cuatro codones cuya frecuencia de uso es inferior al 20 % y un valor CAI de 0.74, lo cual indicaría que la secuencia génica podría ser optimizada para aumentar la expresión. Otras estrategias que podrían ser abordadas para aumentar la expresión de la progaine B serían: la optimización del medio de cultivo, incluyendo el estudio de su producción en forma inactiva (a pH neutro o alcalino) o en su forma activa (a pH ácido) y la expresión de la enzima sin el dominio correspondiente al PSI, que ha demostrado ser un método efectivo para aumentar la producción de la cardosina B producida en *K. lactis* (Almeida et al. 2014; Almeida et al. 2017).

La imposibilidad de mejorar la cepa de producción implica tener que concentrar el sobrenadante del medio de cultivo para poder utilizarlo en la evaluación de la aptitud quesera de la enzima. Las proteasas secretadas por la cepa silvestre de *P. pastoris* al medio de cultivo no tienen actividad coagulante de la leche en las condiciones habituales de elaboración del queso, lo cual implica una gran ventaja para la aplicación industrial de la progaine B, ya que no se requiere un proceso de purificación para su uso, con la reducción del coste de producción que ello implica. La progaine B demostró ser válida para la elaboración de queso fresco a partir de leche cruda de vaca y de queso curado elaborado con leche pasteurizada de oveja. El tipo de coagulante utilizado, progaine B o quimosina comercial (Chymax® Plus, CHR Hansen), tiene influencia en los parámetros de textura, sin embargo, la composición nutricional de los quesos elaborados con

ambos coagulantes es similar. Sorprendentemente, el peso de la masa quesera obtenida con la progaline B fue un 3.14 % superior a la obtenida con la quimosina, normalmente una mayor actividad proteolítica sobre las caseínas debería implicar un menor rendimiento por la pérdida de péptidos en el suero (Agboola et al. 2009; Bruno et al. 2010; Jacob et al. 2011; Mazorra-Manzano et al. 2013b). El coagulante VRen, constituido por el sobrenadante del cultivo de una cepa de *K. lactis* con el gen de la cardosina B, también presenta un mayor rendimiento quesero que la quimosina en quesos elaborados con leche de cabra, oveja y vaca (Almeida et al. 2014). El origen de la leche utilizada y el proceso de elaboración influían, como era de esperar, en los parámetros de textura y nutricionales de los quesos elaborados con la progaline B, así las características del queso fresco de vaca y el madurado de oveja presentaban diferencias significativas. La evaluación global de las propiedades organolépticas realizada por personal no experto de ambos tipos de quesos fue positiva.

El coagulante desarrollado puede ser almacenado utilizando los mismos conservantes y condiciones que otros coagulantes comerciales sin que la progaline B pierda actividad durante un periodo mínimo de seis meses.

El coagulante obtenido podría tener futuro en la industria láctea, ya que originaría masas queseras con características novedosas, ocupando un nicho de mercado diferente al ocupado por los cuajos y coagulantes existentes. El coagulante se produciría en fermentadores a partir de la levadura recombinante y en principio no sería necesario ningún paso de purificación, de manera que se obtendría en grandes cantidades, de una forma fácil, reproducible y con un bajo coste. Para dar el salto del laboratorio al mercado sería necesario obtener una cepa hiperproductora de la enzima y obtener unos parámetros de elaboración de queso adecuados a las características enzimáticas de la progaline B. La mayoría de las investigaciones sobre coagulantes vegetales han estado enfocadas a la búsqueda de un “sustituto” de la

quimosina, en lugar de una “alternativa” o “compañero” de la misma. Así, habría que diseñar un proceso de elaboración adaptado a estas enzimas o contemplar su uso en combinación con el cuajo o coagulantes animales.

La cepa de *P. pastoris* productora de la progaline B ha sido depositada según el Tratado de Budapest, a los fines del procedimiento en materia de patentes, en la Colección Española de Cultivos Tipo con el número CECT 13141 y se ha presentado ante la OEPM una solicitud de patente internacional (PCT/ES2017/070832; título de la invención: “Cepa recombinante, método de producción de proteasas aspárticas de *G. verum* y uso en la industria láctea”). Las proteasas procedentes de *G. verum* figuran en la lista de enzimas para ser evaluadas para su uso en productos alimenticios dentro de la Unión Europea con el número de referencia 2015/54 (Regulation European Commission No 1332/2008).





CONCLUSIONES



6. CONCLUSIONES

De los resultados obtenidos en la presente Tesis Doctoral se han extraído las siguientes conclusiones:

- 1) Los extractos de la planta *Galium verum* tienen actividad coagulante de la leche. Esta actividad es detectada en hojas y flores y debida a la acción de proteasas de tipo aspártico.
- 2) El látex de *Ficus carica* presenta actividad coagulante de la leche. Esta actividad es debida principalmente a la acción de cisteín-proteasas.
- 3) La técnica RACE PCR nos ha permitido obtener la secuencia de dos genes que codifican para dos proteasas aspárticas de *G. verum* (*gv1AP* y *gv2AP*, código Genbank: JQ434478 y JQ434479 respectivamente) y una proteasa aspártica de *F. carica* (*fc1AP*, código Genbank: JQ434480), así como la secuencia parcial de dos genes de esta última especie (*fc2AP* y *fc3AP*, Genbank: JQ434481 y JQ434482, respectivamente).
- 4) El análisis *in silico* de la estructura primaria de las PAs codificadas por los genes aislados indica que presentan los motivos que caracterizan a las PAs vegetales clasificadas como típicas.
- 5) La ubicación de los seis intrones identificados en el gen *gv1AP* está muy conservada entre las proteasas aspárticas típicas, pero no así su longitud. El gen *gv1AP* carece del intrón situado entre los exones 11 y 12 presente en otras especies vegetales.

- 6) La clonación de los genes *gv1AP* y *fc1AP* utilizando vectores tipo pET y el hospedador procariota *E. coli*, permite su expresión en la fracción no soluble. La solubilización de los cuerpos de inclusión con un agente desnaturizante y el proceso de renaturalización mediante el método de dilución rápida no nos permitió la obtención de las proteasas en su forma activa.
- 7) La secuencia codificante del gen *gv2AP* sin el péptido señal se integró en el genoma de la levadura *Pichia pastoris* mediante recombinación homóloga utilizando el vector integrativo pGAPZ α A. Se obtuvo una cepa de *Pichia pastoris* que expresa de manera constitutiva el gen *gv2AP*, secretando la proteasa progaine B al medio de cultivo.
- 8) La progaine B producida por la cepa recombinante no necesita ser activada *in vitro*, se obtiene en su forma activa en el medio de cultivo. La activación ocurre durante el crecimiento de la levadura debido a una bajada de pH del medio durante el crecimiento celular. Tamponar el pH del medio de cultivo nos permite producir la enzima en su forma activa o inactiva.
- 9) El pH, temperatura y termoestabilidad de la progaine B son apropiados para su aplicación en la elaboración de quesos.
- 10) La progaine B tiene actividad coagulante de la leche, el perfil electroforético del corte de la κ -caseína es el mismo que el de otros coagulantes comerciales, pero presenta una mayor actividad proteolítica sobre la α - y β - caseínas.
- 11) El sobrenadante del cultivo de la cepa de *P. pastoris* productora de progaine B concentrado puede ser utilizado

directamente como coagulante para la producción de quesos. Los quesos de vaca elaborados con este coagulante tienen unas características de textura y sensoriales diferentes a las presentadas por los elaborados con la quimosina bovina.







BIBLIOGRAFÍA



7. BIBLIOGRAFÍA

- Abd El-Sal BAE-Y, Ibrahim OAE-H, El-Sayed HAE-R (2017) Purification and characterization of milk clotting enzyme from artichoke (*Cynara cardunculus* L.). Flowers as coagulant on white soft cheese. *Int J Dairy Sci* 12:254–265. doi: 10.3923/ijds.2017.254.265
- Adrio J, Demain A (2014) Microbial enzymes: tools for biotechnological processes. *Biomolecules* 4:117–139. doi: 10.3390/biom4010117
- Agboola S, Chen S, Zhao J (2004) Formation of bitter peptides during ripening of ovine milk cheese made with different coagulants. *Lait* 84:567–578. doi: 10.1051/lait:2004032
- Agboola SO, Chan HH, Zhao J, Rehman A (2009) Can the use of Australian cardoon (*Cynara cardunculus* L.) coagulant overcome the quality problems associated with cheese made from ultrafiltered milk?. *LWT-Food Sci Technol* 42:1352–1359. doi: 10.1016/j.lwt.2009.03.015
- Ageitos JM, Vallejo JA, Poza M, Villa TG (2006) Fluorescein thiocarbamoyl-kappa-casein assay for the specific testing of milk-clotting proteases. *J Dairy Sci* 89:3770–3777. doi: 10.3168/jds.S0022-0302(06)72418-3
- Ahmad M, Hirz M, Pichler H, Schwab H (2014) Protein expression in *Pichia pastoris*: Recent achievements and perspectives for heterologous protein production. *Appl Microbiol Biotechnol* 98:5301–5317. doi: 10.1007/s00253-014-5732-5
- Ahmed IAM, Babiker EE, Mori N (2010) pH stability and influence of salts on activity of a milk-clotting enzyme from *Solanum dubium* seeds and its enzymatic action on bovine caseins. *LWT Food Sci Technol* 43:759–764. doi: 10.1016/j.lwt.2009.12.011

- Ahmed SA, Helmy WA (2012) Comparative evaluation of *Bacillus licheniformis* 5A5 and *Aloe variegata* milk-clotting enzymes. Brazilian J Chem Eng 29:69–76. doi: 10.1590/S0104-66322012000100008
- Ahn VE, Leyko P, Alattia J-R, Chen L, Privé GG (2006) Crystal structures of saposins A and C. Protein Sci 15:1849–57. doi: 10.1110/ps.062256606
- Akinloye AM, Adewumi OO (2014) Effects of local coagulants on the yield of cheese using cow and sheep milk. Int J Dev Sustain Int J Dev Sustain 33:2168–8662.
- Akpınar O, Penner MH (2002). Peptidase Activity Assays Using Protein Substrates. In: Current Protocols in Food Analytical Chemistry. Wiley, New Jersey, C 2.2.1-C.2.2.2.
- Alirezaei M, Aminlari M, Gheisari HR, Tavana M (2011) Actinidin: A promising milk coagulating enzyme. Eur J Food Res Rev 1:43–51.
- Almeida CM, Gomes D, Faro C, Simões I (2014) Engineering a cardosin B-derived rennet for sheep and goat cheese manufacture. Appl Microbiol Biotechnol 99:269–281. doi: 10.1007/s00253-014-5902-5
- Almeida CM, Manso JA, Figueiredo AC, Antunes L, Cruz R, Manadas B, Bur D, Pereira PJB, Faro C, Simões I (2017) Functional and structural characterization of synthetic cardosin B-derived rennet. Appl Microbiol Biotechnol 101:6951–6968. doi: 10.1007/s00253-017-8445-8
- Almeida CM, Pereira C, da Costa DS, Pereira S, Pissarra J, Simões I, Faro C, Feil R, Lunn J, Brust H, Freeling M (2012) Chlpsin, a chloroplastidial aspartic proteinase from the green algae *Chlamydomonas reinhardtii*. Planta 236:283–296. doi: 10.1007/s00425-012-1605-2
- Allison RG, Senti FR (1983) A perspective on the application of the

- Atwater system of food energy assessment. Bethesda: Life Sciences Research Office, Federation of American Societies for Experimental Biology.
- An CI, Fukusaki E, Kobayashi A (2002) Aspartic proteinases are expressed in pitchers of the carnivorous plant *Nepenthes alata* Blanco. *Planta* 214:661–667. doi: 10.1007/s004250100665
- Anderson DH, Sawaya MR, Cascio D, Ernst W, Modlin R, Krensky A, Eisenberg D (2003) Granulysin crystal structure and a structure-derived lytic mechanism. *J Mol Biol* 325:355–65.
- Anusha R, Singh MK, Bindhu OS (2013) Screening of latex producing plants for their milk clotting activity. *Res J Pharm Biol Chem Sci* 4:757–761.
- Anusha R, Singh MK, Bindhu OS (2014) Characterisation of potential milk coagulants from *Calotropis gigantea* plant parts and their hydrolytic pattern of bovine casein. *Eur Food Res Technol* 238:997–1006. doi: 10.1007/s00217-014-2177-0
- AOAC (2002) Association of Official Analytical Chemists. Official methods of analysis (17th ed). Gaithersburg: AOAC.
- Aquilanti L, Babini V, Santarelli S, Osimani A, Petruzzelli A, Clementi F (2011) Bacterial dynamics in a raw cow's milk Caciotta cheese manufactured with aqueous extract of *Cynara cardunculus* dried flowers. *Lett Appl Microbiol* 52:651–659. doi: 10.1111/j.1472-765X.2011.03053.x
- Araya-Garay JM, Feijoo-Siota L, Rosa-Dos-Santos F, Veiga-Crespo P, Villa TG (2012) Construction of new *Pichia pastoris* X-33 strains for production of lycopene and β -carotene. *Appl Microbiol Biotechnol*. doi: 10.1007/s00253-011-3764-7
- Arima K, Yu J, Iwasaki S (1970) Milk-clotting enzyme from *Mucor pusillus* var. *Lindt*. *Meth. Enzymol* 19: 446–459.
- Asakura T, Matsumoto I, Funaki J, Arai S, Abe K (2000) The plant

- aspartic proteinase-specific polypeptide insert is not directly related to the activity of oryzasin 1. *Eur J Biochem* 267:5115–5122. doi: 10.1046/j.1432-1327.2000.01582.x
- Asakura T, Watanabe H, Abe K, Arai S (1995) Rice aspartic proteinase, oryzasin, expressed during seed ripening and germination, has a gene organization distinct from those of animal and microbial aspartic proteinases. *Eur J Biochem* 232:77–83. doi: 10.1111/j.1432-1033.1995.tb20783.x
- Asakura T, Watanabe H, Abe K, Arai S (1997) Oryzasin as an aspartic proteinase occurring in rice seeds: purification, characterization, and application to milk clotting. *J Agric Food Chem* 45:1070–1075. doi: 10.1021/jf960582x
- Athauda SB, Matsumoto K, Rajapakshe S, Kuribayashi M, Kojima M, Kubomura-Yoshida N, Iwamatsu A, Shibata C, Inoue H, Takahashi K (2004) Enzymic and structural characterization of nepenthesin, a unique member of a novel subfamily of aspartic proteinases. *Biochem J* 381:295–306. doi: 10.1042/BJ20031575
- Aworh OC, Muller HG (1987) Cheese-making properties of vegetable rennet from sodom apple (*Calotropis procera*). *Food Chem* 26:71–79. doi: 10.1016/0308-8146(87)90168-3
- Badgujar SB, Mahajan RT (2014) Nivulian-II a new milk clotting cysteine protease of *Euphorbia nivulia* latex. *Int J Biol Macromol* 70:391–398. doi: 10.1016/j.ijbiomac.2014.07.022
- Barbosa, M., Corradini, C. and Battistoni, B. (1981) Cheese-making experiments carried out on some Italian cheeses with vegetable rennet from cardo (*Cynara cardunculus* L.). *Sci Tecn Latt Cas* 32: 203-221.
- Barret AJ (1997) Nomenclature Committee of the International Union of Biochemistry and Molecular Biology (NC-IUBMB). *Enzyme Nomenclature. Recommendations 1992. Supplement 4: corrections and additions* (1997). *Eur J Biochem* 250:1-6.

- Barros RM, Ferreira CA, Silva S V, Malcata FX (2001) Quantitative studies on the enzymatic hydrolysis of milk proteins brought about by cardosins precipitated by ammonium sulfate. *Enzyme Microb Technol* 29:541–547. doi: 10.1016/S0141-0229(01)00431-8
- Beka RG, Krier F, Botquin M, Guiana VD, Donn P, Libouga DG, Mbofung CM, Dimitrov K, Slomianny M-C, Guillochon D, Vercaigne-Marko D (2014) Characterisation of a milk-clotting extract from *Balanites aegyptiaca* fruit pulp. *Int Dairy J* 34:25–31. doi: 10.1016/j.idairyj.2013.06.013
- Belozersky MA, Sarbakanova ST, Dunaevsky YE (1989) Aspartic proteinase from wheat seeds: isolation, properties and action on gliadin. *Planta* 177:321–326. doi: 10.1007/BF00403589
- Ben Amira A, Besbes S, Attia H, Blecker C (2017a) Milk-clotting properties of plant rennets and their enzymatic, rheological and sensory role in cheese making: A review. *Int J Food Prop* 20:S76-S93 doi: 10.1080/10942912.2017.1289959
- Ben Amira A, Makhlouf I, Flaviu Petrut R, Francis F, Bauwens J, Attia H, Besbes S, Blecker C (2017b) Effect of extraction pH on techno-functional properties of crude extracts from wild cardoon (*Cynara cardunculus* L.) flowers. *Food Chem* 225:258–266. doi: 10.1016/J.FOODCHEM.2017.01.040
- Bernstein NK, James MN (1999) Novel ways to prevent proteolysis — prophytepsin and proplasmepsin II. *Curr Opin Struct Biol* 9:690–695. doi: 10.1016/S0959-440X(99)00030-5
- Berridge NJ (1952) Some observations on the determination of the activity of rennet. *Analyst* 77:57-62.
- Bi X, Khush GS, Bennett J (2005) The rice nucellin gene ortholog *OsAsp1* encodes an active aspartic protease without a plant-specific insert and is strongly expressed in early embryo. *Plant Cell Physiol* 46:87–98. doi: 10.1093/pcp/pci002

- Bleukx W, Torrekensl S, Van Leuvenl F, Delcour JA (1998) Purification, properties and N-terminal amino acid sequence of a wheat gluten aspartic proteinase. *J Cereal Sci* 28:223–232.
- Bliven S, Prlić A (2012) Circular permutation in proteins. *PLoS Comput Biol*. 8:e1002445 doi: 10.1371/journal.pcbi.1002445
- Böjthe-Horváth K, Hetényi F, Kocsis Á, Szabó L, Varga-Balázs M, Máthé I Jr and Tétényi P (1980) Iridoid glycosides from *Galium verum*. *Phytochemistry* 21: 2917-2919.
- Borisov MI, Zaitsev KV, Zaitsev VG (1971) The chemical composition of *Galium verum*. *Chem Nat Compd* 7: 511.
- Bornaz S, Guizani N, Fellah N, Sahli A, Slama M Ben, Attia H (2010) Effect of plant originated coagulants and chymosin on ovine milk coagulation. *Int J Food Prop* 13:10–22. doi: 10.1080/10942910802144238
- Bourgeois J, Malek L (1991) Purification and characterization of an aspartyl proteinase from dry jack pine seeds. *Seed Sci Res* 1:139–147. doi: 10.1017/S0960258500000817
- Bradford MM (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem* 72:248-54.
- Bretthauer RK, Castellino FJ (1999) Glycosylation of *Pichia pastoris*-derived proteins. *Biotechnol Appl Biochem* 30(Pt 3):193–200. doi: 10.1111/j.1470-8744.1999.tb00770.x
- Brodelius M, Hiraiwa M, Marttila S, Karadaghi S Al, Picaud S, Brodelius PE (2005) Immunolocalization of the saposin-like insert of plant aspartic proteinases exhibiting saposin C activity. Expression in young flower tissues and in barley seeds. *Physiol Plant* 125:405–418. doi: 10.1111/j.1399-3054.2005.00576.x
- Brown JWS, Simpson CG (1998) Splice site selection in plant pre-mRNA splicing. 77–95.

- Bruhn H (2005) A short guided tour through functional and structural features of saposin-like proteins. *Biochem J* 389:249–57. doi: 10.1042/BJ20050051
- Bruno MA, Lazza CM, Errasti ME, López LMI, Caffini NO, Pardo MF (2010) Milk clotting and proteolytic activity of an enzyme preparation from *Bromelia hieronymi* fruits. *LWT-Food Sci Technol* 43:695–701. doi: 10.1016/j.lwt.2009.12.003
- Brutti CB, Pardo MF, Caffini NO, Natalucci CL (2012) *Onopordum acanthium* L. (Asteraceae) flowers as coagulating agent for cheesemaking. *LWT-Food Sci Technol* 45:172–179. doi: 10.1016/j.lwt.2011.09.001
- Bryksa BC, Bhaumik P, Magracheva E, De Moura DC, Kurylowicz M, Zdanov A, Dutcher JR, Wlodawer A, Yada RY (2011) Structure and mechanism of the saposin-like domain of a plant aspartic protease. *J Biol Chem* 286:28265–28275. doi: 10.1074/jbc.M111.252619
- Bryksa BC, Grahame DA, Yada RY (2017) Comparative structure-function characterization of the saposin-like domains from potato, barley, cardoon and *Arabidopsis* aspartic proteases. *Biochim Biophys Acta* 1859:1008–1018. doi: 10.1016/j.bbamem.2017.02.007
- Buch F, Kaman WE, Bikker FJ, Yilamujiang A, Mithöfer A (2015) Nepenthesin protease activity indicates digestive fluid dynamics in carnivorous nepenthes plants. *PLoS One* 10:e0118853. doi: 10.1371/journal.pone.0118853
- Campos R, Guerra R, Aguilar M, Ventura O, Camacho L (1990) Chemical characterization of proteases extracted from wild thistle (*Cynara cardunculus*). *Food Chem* 35:89–97. doi: 10.1016/0308-8146(90)90023-W
- Castanheira P, Samyn B, Sergeant K, Clemente JC, Dunn BM, Pires E, Van Beeumen J, Faro C (2005) Activation, proteolytic

- processing, and peptide specificity of recombinant cardosin A. *J Biol Chem* 280:13047–13054. doi: 10.1074/jbc.M412076200
- Cereghino JL, Cregg JM (2000) Heterologous protein expression in the methylotrophic yeast *Pichia pastoris*. *FEMS Microbiol Rev* 24:45–66.
- Chávez-Garay DR, Gutiérrez-Méndez N, Valenzuela-Soto ME, García-Triana A (2016) Partial characterization of a plant coagulant obtained from the berries of *Solanum elaeagnifolium*. *J food* 14:200–205. doi: 10.1080/19476337.2015.1080763
- Chazarra S, Sidrach L, López-Molina D, Rodríguez-López JN (2007) Characterization of the milk-clotting properties of extracts from artichoke (*Cynara scolymus*, L.) flowers. *Int Dairy J* 17:1393–1400. doi: 10.1016/j.idairyj.2007.04.010
- Chen F, Foolad MR (1997) Molecular organization of a gene in barley which encodes a protein similar to aspartic protease and its specific expression in nucellar cells during degeneration. *Plant Mol Biol* 35:821–31.
- Chen HJ, Huang YH, Huang GJ, Huang SS, Chow TJ, Lin YH (2015) Sweet potato SPAP1 is a typical aspartic protease and participates in ethephon-mediated leaf senescence. *J Plant Physiol* 180:1–17. doi: 10.1016/j.jplph.2015.03.009
- Chen J, Ding J, Ouyang Y, Du H, Yang J, Cheng K, Zhao J, Qiu S, Zhang X, Yao J, Liu K, Wang L, Xu C, Li X, Xue Y, Xia M, Ji Q, Lu J, Xu M, Zhang Q (2008) A triallelic system of S5 is a major regulator of the reproductive barrier and compatibility of indica-japonica hybrids in rice. *Proc Natl Acad Sci U S A* 105:11436–41. doi: 10.1073/pnas.0804761105
- Chen J, Ouyang Y, Wang L, Xie W, Zhang Q (2009) Aspartic proteases gene family in rice: Gene structure and expression, predicted protein features and phylogenetic relation. *Gene* 442:108–118. doi: 10.1016/j.gene.2009.04.021

- Chen S, Agboola S, Zhao J (2003a) Use of Australian cardoon extract in the manufacture of ovine milk cheese - A comparison with commercial rennet preparations. *Int J Food Sci Technol* 38:799–807. doi: 10.1046/j.1365-2621.2003.00735.x
- Chen S, Zhao J, Agboola S (2003b) Isolation and partial characterization of rennet-like proteases from Australian Cardoon (*Cynara cardunculus* L.). *J Agric Food Chem* 51:3127–3134. doi: 10.1021/jf025842r
- Chen X, Pfeil JE, Gal S (2002) The three typical aspartic proteinase genes of *Arabidopsis thaliana* are differentially expressed. *Eur J Biochem* 269:4675–4684. doi: 10.1046/j.1432-1033.2002.03168.x
- Chenchik A, Moqadam F, Siebert P (1996) A new method for full-length cDNA cloning by PCR. In: Krieg P (ed) *A laboratory guide to RNA: isolation, analysis, and synthesis*. Wiley-Liss, New York, pp 273–321.
- Chitpintyol S, Crabbe MJC (1998) Chymosin and aspartic proteinases. *Food Chem* 61:395–418.
- Chomé P, López Corrales M, Torralvo MG et al. (2011) Variedades de Higuera. Descripción y registro de variedades. Ministerio de Medio Ambiente y Medio Rural y Marino. ISBN: 84-491-1103-7.
- Cimino CV, Colombo ML, Liggieri C, Bruno M, Vairo-Cavalli S (2015) Partial molecular characterization of *Arctium minus* aspartylendopeptidase and preparation of bioactive peptides by whey protein hydrolysis. *J Med Food* 18:856–864. doi: 10.1089/jmf.2014.0101
- Contour-Ansel D, Torres-Franklin ML, Zuily-Fodil Y, de Carvalho MHC (2010) An aspartic acid protease from common bean is expressed “on call” during water stress and early recovery. *J Plant Physiol* 167:1606–1612. doi: 10.1016/j.jplph.2010.06.018

- Cordeiro M, Lowther T, Dunn BM, Guruprasad K, Blundell T, Pais MS, Brodelius PE (1998) Substrate specificity and molecular modelling of aspartic proteinases (cyprosin) from flowers of *Cynara cardunculus* subsp. *flavescens* cv. *cardoon*. *Adv Exp Med Biol* 436:473-9.
- Cordeiro MC, Jacob E, Puhán Z, S Pais M, Brodelius P (1992) Milk Clotting and proteolytic activities of purified cynarases from *Cynara cardunculus* L.; a comparison to chymosin. *Milchwissenschaft* 47:683–687.
- Cordeiro MC, Pais MS, Brodelius PE (1994a) Tissue-specific expression of multiple forms of cyprosin (aspartic proteinase) in flowers of *Cynara cardunculus*. *Physiol Plant* 92:645–653. doi: 10.1111/j.1399-3054.1994.tb03035.x
- Cordeiro MC, Xue Z-T, Pietrzak M, Salomé Pais M, Brodelius PE (1994b) Isolation and characterization of a cDNA from flowers of *Cynara cardunculus* encoding cyprosin (an aspartic proteinase) and its use to study the organ-specific expression of cyprosin. *Plant Mol Biol* 24:733–741. doi: 10.1007/BF00029855
- Corrons MA, Bertucci JI, Liggieri CS, López LMI, Bruno MA (2012) Milk clotting activity and production of bioactive peptides from whey using *Maclura pomifera* proteases. *LWT-Food Sci Technol* 47:103–109. doi: 10.1016/J.LWT.2011.12.028
- Costa J, Ashford DA, Nimtz M, Bento I, Frazão C, Esteves CL, Faro CJ, Kervinen J, Pires E, Veríssimo P, Wlodawer A, Carrondo MA (1997) The glycosylation of the aspartic proteinases from barley (*Hordeum vulgare* L.) and cardoon (*Cynara cardunculus* L.). *Eur J Biochem* 243:695–700.
- Crabbe MJC (2004) Rennets: General and molecular aspects. In: Fox PF, McSweeney PLH, Cogan TM, Guinee TP (eds) *Cheese: chemistry, physics and microbiology*. 3rd edn, Elsevier Academic Press, London pp 37–68.

- Cregg JM (2007) *Pichia* Protocols. Methods in Molecular Biology. Humana Press, Totowa, New Jersey.
- Cregg JM, Tolstorukov I, Kusari A, Sunga J, Madden K, Chappell T (2009) Expression in the Yeast *Pichia pastoris*. Methods Enzymol 463:169-189.
- Cruz de Carvalho MH, Lameta D, Macauley R, Gareil M, Maarouf E, Thi P, Fodil Z, de Carvalho MHC, d'Arcy Lameta A, Macauley HR, Maarouf H El, Thi ATP, Fodil YZ (2001) Aspartic protease in leaves of common bean (*Phaseolus vulgaris* L.) and cowpea (*Vigna unguiculata* L. Walp): enzymatic activity, gene expression and relation to drought susceptibility. Febs Lett 492:242–246.
- Cruz de Carvalho MH, Pham-Thi AT, Gareil M, d'Arcy-Lameta A, Fodil YZ (2004) Isolation and characterization of an aspartic proteinase gene from cowpea (*Vigna unguiculata* L. Walp.). J Plant Physiol 161:971–976. doi: 10.1016/j.jplph.2003.12.007
- Curto P, Lufrano D, Pinto C, Custodio V, Gomes AC, Trejo SA, Bakas L, Vairo-Cavalli S, Faro C, Simoes I (2014) Establishing the yeast *Kluyveromyces lactis* as an expression host for production of the saposin-like domain of the aspartic protease cirsin. Appl Environ Microbiol 80:86–96. doi: 10.1128/AEM.03151-13
- D 'hondt K, Stack S, Gutteridge S, Vandekerckhove J, Krebbers E, Gal S (1997) Aspartic proteinase genes in the Brassicaceae *Arabidopsis thaliana* and *Brassica napus*. Plant Mol Biol 33:187–192.
- da Costa DS, Pereira S, Moore I, Pissarra J (2010) Dissecting cardosin B trafficking pathways in heterologous systems. Planta 232:1517–1530. doi: 10.1007/s00425-010-1276-9
- Dalgleish DG (1993) The Enzymatic coagulation of milk. In: Cheese: Chemistry, Physics and Microbiology. Springer US, Boston, MA,

pp 69–100.

- Dalgleish DG, Corredig M (2012) The Structure of the casein micelle of milk and its changes during processing. *Annu Rev Food Sci Technol* 3:449–467. doi: 10.1146/annurev-food-022811-101214
- Daly R, Hearn MT (2005) Expression of heterologous proteins in *Pichia pastoris*: a useful experimental tool in protein engineering and production. *J Mol Recognit* 18:119–138. doi: 10.1002/jmr.687
- Damasceno LM, Huang CJ, Batt CA (2012) Protein secretion in *Pichia pastoris* and advances in protein production. *Appl Microbiol Biotechnol* 93:31–39. doi: 10.1007/s00253-011-3654-z
- Darabi M, Seddigh S (2015) Bioinformatic characterization of aspartic protease (AP) enzyme in seed plants. *Plant Syst Evol* 301: 2399–2417.
- Davies DR (1990) The structure and function of the aspartic proteinases. *Annu Rev Biophys Biophys Chem* 19:189–215. doi: 10.1146/annurev.bb.19.060190.001201
- de Sá FV, Barbosa M (1972) Cheese-making with a vegetable rennet from cardo (*Cynara cardunculus*). *J Dairy Res* 39:335–343. doi: 10.1017/S0022029900014163
- Demir Y, Güngör A, Sarikaya Öztürk SB, Demir N (2007) The purification of protease from cowslip (*Primula veris*) and its use in food processing. *J Food Process Preserv* 31:559–570. doi: 10.1111/j.1745-4549.2007.00145.x
- Demir Y, Alayli A, Yildirim, Demir N (2005) Identification of protease from *Euphorbia amygdaloides* latex and its use in cheese production. *Prep Biochem Biotechnol* 35:291–299. doi: 10.1080/10826060500218107
- Devaraj KB, Gowda LR, Prakash V (2008) An unusual thermostable aspartic protease from the latex of *Ficus racemosa* (L.).

Phytochemistry 69:647–655.

- Doi E, Shibata D, Matoba T, Yonezawa D (1980) Characterization of pepstatin-sensitive acid protease in resting rice seeds. *Agr Biol Chem* 44:741–747
- Dombrowski JE, Schroeder MR, Bednarek SY, Raikhel NV (1993) Determination of the functional elements within the vacuolar targeting signal of barley lectin. *Plant Cell* 5:587–596.
- Domingos A, Cardoso PC, Xue ZT, Clemente A, Brodelius PE, Pais MS (2000) Purification, cloning and autoproteolytic processing of an aspartic proteinase from *Centaurea calcitrapa*. *Eur J Biochem* 267:6824–6831. doi: 10.1046/j.1432-1327.2000.01780.x
- Domingos A, Xue Z-T, Guruprasad K, Clemente A, Blundell T, Pais MS, Brodelius PE (1998) An aspartic proteinase from flowers of *Centaurea calcitrapa*. In: James MNG (ed) *Aspartic proteinases*. Springer, Boston, MA, pp 465–472.
- Duarte AR, Duarte DMR, Moreira KA, Cavalcanti MTH, Lima-Filho JL de, Porto ALF (2009) *Jacaratia corumbensis* O. Kuntze a new vegetable source for milk-clotting enzymes. *Brazilian Arch Biol Technol* 52:1–9. doi: 10.1590/S1516-89132009000100001
- Duarte P, Pissarra J, Moore I (2008) Processing and trafficking of a single isoform of the aspartic proteinase cardosin a on the vacuolar pathway. *Planta* 227:1255–1268. doi: 10.1007/s00425-008-0697-1
- Dunn BM (2002) Structure and mechanism of the pepsin-like family of aspartic peptidases. *Chem Rev* 102:4431–58.
- Edgar RC (2004) MUSCLE: Multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* 32:1792–1797. doi: 10.1093/nar/gkh340
- Egas C, Lavoura N, Resende R, Brito RMM, Pires E, Pedroso De

- Lima MC, Faro C (2000) The saposin-like domain of the plant aspartic proteinase precursor is a potent inducer of vesicle leakage. *J Biol Chem* 275:38190–38196. doi: 10.1074/jbc.M006093200
- Egito AS, Girardet JM, Laguna LE, Poirson C, Mollé D, Miclo L, Humbert G, Gaillard JL (2007) Milk-clotting activity of enzyme extracts from sunflower and albizia seeds and specific hydrolysis of bovine κ -casein. *Int Dairy J* 17:816–825. doi: 10.1016/j.idairyj.2006.09.012
- El-Kholy AM (2015) Ras cheese making with vegetable coagulant -a comparison with calf rennet. *World J Dairy Food Sci* 10:82–89. doi: 10.5829/idosi.wjdfs.2015.10.1.95110
- Espinoza-Molina JA, Acosta-Muñiz CH, Sepulveda DR, Zamudio-Flores PB, Rios-Velasco C (2016) Codon optimization of the “*Bos taurus* chymosin” gene for the production of recombinant chymosin in *Pichia pastoris*. *Mol Biotechnol* 58:657–664. doi: 10.1007/s12033-016-9965-7
- Esposito M, Di Pierro P, Dejonghe W, Mariniello L, Porta R, Pierro D, Dejonghe W, Mariniello L, Porta R (2016) Enzymatic milk clotting activity in artichoke (*Cynara scolymus*) leaves and alpine thistle (*Carduus defloratus*) flowers. Immobilization of alpine thistle aspartic protease. *Food Chem* 204:115–121. doi: 10.1016/j.foodchem.2016.02.060
- Exterkate FA, Lagerwerf FM, Haverkamp J, Van Schalkwijk S (1997) The selectivity of chymosin action on α s1- and β -caseins in solution is modulated in cheese. *Int Dairy J* 7:41–54.
- Faccia M, Picariello G, Trani A, Loizzo P, Gambacorta G, Lamacchia C, Di Luccia A (2012) Proteolysis of Cacioricotta cheese made from goat milk coagulated with caprifig (*Ficus carica sylvestris*) or calf rennet. *Eur Food Res Technol* 234:527–533. doi: 10.1007/s00217-012-1668-0

- Faro C, Gal S (2005) Aspartic proteinase content of the *Arabidopsis* genome. *Curr Protein Pept Sci* 6:493–500. doi: 10.2174/138920305774933268
- Faro C, Ramalho-Santos M, Vieira M, Mendes A, Simoes I, Andrade R, Veríssimo P, Lin XL, Tang J, Pires E (1999) Cloning and characterization of cDNA encoding cardosin A, all RGD-containing plant aspartic proteinase. *J Biol Chem* 274:28724–28729. doi: 10.1074/jbc.274.40.28724
- Feijoo-Siota L, Blasco L, Rodríguez-Rama JL, Barros-Velázquez J, Miguel TD, Sánchez-Pérez A, Villa TG (2014) Recent patents on microbial proteases for the dairy industry. *Recent Adv DNA Gene Seq* 8:44-55.
- Feijoo-Siota L, Villa TG (2011) Native and Biotechnologically Engineered plant proteases with industrial applications. *Food Bioprocess Technol* 4:1066–1088. doi: 10.1007/s11947-010-0431-4
- Felix da Silva D, Matumoto-Pintro PT, Bazinet L, Couillard C, Britten M (2015) Effect of commercial grape extracts on the cheese-making properties of milk. *J Dairy Sci* 98:1552–1562. doi: 10.3168/jds.2014-8796
- Felsenstein J (1985) Confidence limits on phylogenies: an approach using the bootstrap. *Evolution* (NY) 39:783–791. doi: 10.1111/j.1558-5646.1985.tb00420.x
- Fernandez-Salguero J, Prados F, Calixto F, Vioque M, Sampaio P, Tejada L (2003) Use of recombinant cyprosin in the manufacture of ewe's milk cheese. *J Agric Food Chem* 51:7426–7430. doi: 10.1021/jf034573h
- Fernández-Salguero J, Tejada L, Gómez R (2002) Use of powdered vegetable coagulant in the manufacture of ewe's milk cheeses. *J Sci Food Agric* 82:464–468. doi: 10.1002/jsfa.1066
- Fernández-Salguero J, Sánchez E, Gómez R, Mata C, Vioque M,

- Tejada L (1999) A preliminary study of microbiological quality of cardoons of the genus *Cynara* L. used in manufacture of traditional cheeses. *Michwissenschaft*. 54: 688-690.
- Finn RD, Coghill P, Eberhardt RY, Eddy SR, Mistry J, Mitchell AL, Potter SC, Punta M, Qureshi M, Sangrador-Vegas A, Salazar GA, Tate J, Bateman A (2016) The Pfam protein families database: towards a more sustainable future. *Nucleic Acids Res* 44:D279–D285. doi: 10.1093/nar/gkv1344
- Flamm EL (1991) How FDA approved chymosin: A case history. *Bio/Technology* 9:349–351. doi: 10.1038/nbt0491-349
- Fox PF, Guinee TP, Cogan TM, McSweeney PLH (2017) *Fundamentals of Cheese Science*. Springer US, Boston, MA.
- Fox PF, McSweeney PLH (1997) Rennets: their role in milk coagulation and cheese ripening. In: Law BA (ed) *Microbiology and Biochemistry of Cheese and Fermented Milk*. Springer US, Boston, MA, pp 1–49.
- Francisco-José D, Joaquín R-P, José G-C, Ramírez R, Isidro R (2010) Proteolysis and texture changes of a Spanish soft cheese (“Torta del Casar”) manufactured with raw ewe milk and vegetable rennet during ripening. *Int J Food Sci Technol* 45:512–519. doi: 10.1111/j.1365-2621.2009.02157.x
- Frazão C, Bento I, Costa J, Soares CM, Veríssimo P, Faro C, Pires E, Cooper J, Carrondo MA (1999) Crystal structure of cardosin A, a glycosylated and Arg-Gly-Asp-containing aspartic proteinase from the flowers of *Cynara cardunculus* L. *J Biol Chem* 274:27694–27701. doi: 10.1074/jbc.274.39.27694
- Freitas AC, Malcata FX (1996) Influence of milk type, coagulant, salting procedure and ripening time on the final characteristics of Picante cheese. *Int Dairy J* 6:1099–1116. doi: 10.1016/S0958-6946(96)00035-0
- Freitas CDT, Leite HB, Oliveira JPB, Amaral JL, Egito AS, Vairo-

- Cavalli S, Lobo MDP, Monteiro-Moreira ACO, Ramos M V. (2016) Insights into milk-clotting activity of latex peptidases from *Calotropis procera* and *Cryptostegia grandiflora*. Food Res Int 87:50–59. doi: 10.1016/j.foodres.2016.06.020
- Frey ME, D'Ippolito S, Pepe A, Daleo GR, Guevara MG (2018) Transgenic expression of plant-specific insert of potato aspartic proteases (StAP-PSI) confers enhanced resistance to *Botrytis cinerea* in *Arabidopsis thaliana*. Phytochemistry 149:1-11. doi: 10.1016/j.phytochem.2018.02.004
- Frith GJT, Peoples MB, Dalling MJ (1978) Proteolytic enzymes in green wheat-leaves III. Inactivation of acid proteinase II by diazoacetyl-DL-norleucine methyl ester and 1,2-epoxy-3-(p-nitrophenoxy)-propane. Plant Cell Physiol 19:819–824. doi: 10.1093/oxfordjournals.pcp.a075656
- Fujikura Y, Karssen CM (1995) Molecular studies on osmoprimed seeds of cauliflower: a partial amino acid sequence of a vigour-related protein and osmopriming-enhanced expression of putative aspartic protease. Seed Sci Res 5:177–181. doi: 10.1017/S0960258500002804
- Gagaoua M, Ziane F, Nait Rabah S, Boucherba N, Ait Kaki El-Hadef El-Okki A, Bouanane-Darenfed A, Hafid K (2017) Three phase partitioning, a scalable method for the purification and recovery of cucumisin, a milk-clotting enzyme, from the juice of *Cucumis melo* var. *reticulatus*. Int J Biol Macromol 102:515–525. doi: 10.1016/J.IJBIOMAC.2017.04.060
- Galan E, Cabezas L, Fernandez-Salguero J (2012) Proteolysis, microbiology and sensory properties of ewes' milk cheese produced with plant coagulant from cardoon *Cynara cardunculus*, calf rennet or a mixture thereof. Int Dairy J 25:92–96. doi: 10.1016/j.idairyj.2012.02.001
- Galan E, Prados F, Pino A, Tejada L, Fernandez-Salguero J, Galán E, Prados F, Pino A, Tejada L, Fernández-Salguero J (2008)

- Influence of different amounts of vegetable coagulant from cardoon *Cynara cardunculus* and calf rennet on the proteolysis and sensory characteristics of cheeses made with sheep milk. *Int Dairy J* 18:93–98. doi: 10.1016/j.idairyj.2007.06.003
- Galleschi L, Felicioli F (1994) Purification, characterization and activation by anions of an aspartic proteinase isolated from bran of soft wheat. *Plant Sci* 98:15–24. doi: 10.1016/0168-9452(94)90143-0
- Gama Salgado JA, Kangwa M, Fernandez-Lahore M (2013) Cloning and expression of an active aspartic proteinase from *Mucor circinelloides* in *Pichia pastoris*. *BMC Microbiol* 13:250. doi: 10.1186/1471-2180-13-250
- García V, Rovira S, Teruel R, Boutoia K, Rodríguez J, Roa I, López MB (2012) Effect of vegetable coagulant, microbial coagulant and calf rennet on physicochemical, proteolysis, sensory and texture profiles of fresh goats cheese. *Dairy Sci Technol* 92:691–707. doi: 10.1007/s13594-012-0086-1
- Ge X, Dietrich C, Matsuno M, Li G, Berg H, Xia Y (2005) An Arabidopsis aspartic protease functions as an anti-cell-death component in reproduction and embryogenesis. *EMBO Rep* 6:282–8. doi: 10.1038/sj.embor.7400357
- Glathe S, Kervinen J, Nimtz M, Li GH, Tobin GJ, Copeland TD, Ashford DA, Wlodawer A, Costa J (1998) Transport and activation of the vacuolar aspartic proteinase phytepsin in barley (*Hordeum vulgare* L.). *J Biol Chem* 273:31230–31236. doi: 10.1074/jbc.273.47.31230
- Goettig P (2016) Effects of glycosylation on the enzymatic activity and mechanisms of proteases. *Int J Mol Sci*. doi: 10.3390/ijms17121969
- Gómez R, Sánchez E, Vioque M, Ferreira J, Tejada L, Fernández-Salguero J (2001) Microbiological characteristics of ewes' milk

- cheese manufactured using aqueous extracts of flowers from various species of cardoon *Cynara* L. *Milchwissenschaft* 56:16-19.
- González-Rodríguez AM, Grajal-Martín MJ (2011) “Higueras de Canarias. Caracterización morfológica de variedades”. Instituto Canario de Investigaciones Agrarias. Gobierno de Canarias.
- Grappin R, Rank TC, Olson NF (1985) Primary proteolysis of cheese proteins during ripening. A review. *J Dairy Sci* 68:531–540. doi: 10.3168/jds.S0022-0302(85)80855-9
- Grieve M (1971) *A Modern Herbal*. Vol I:91-92.
- Grinna LS, Tschopp JF (1989) Size distribution and general structural features of N-linked oligosaccharides from the methylotrophic yeast, *Pichia pastoris*. *Yeast* 5:107–115. doi: 10.1002/yea.320050206
- Gromova I, Celis JE (2006) Protein detection in gels by silver staining: a procedure compatible with mass spectrometry. In Celis JE, Carter N, Hunter T, Simons K, Small JV, Shotton D (eds.) *Cell Biology a Laboratory Handbook*, pp. 219–223, Elsevier Science, New York, NY, USA, 3rd edition.
- Grozdanovic MM, Burazer L, Gavrovic-Jankulovic M (2013) Kiwifruit (*Actinidia deliciosa*) extract shows potential as a low-cost and efficient milk-clotting agent. *Int Dairy J* 32:46–52. doi: 10.1016/J.IDAIRYJ.2013.03.001
- Guevara MG, Almeida C, Mendieta JR, Faro CJ, Veríssimo P, Pires EV, Daleo GR (2005) Molecular cloning of a potato leaf cDNA encoding an aspartic protease (StAsp) and its expression after *P. infestans* infection. *Plant Physiol Biochem* 43:882–889. doi: 10.1016/j.plaphy.2005.07.004
- Guevara MG, Daleo GR, Oliva CR (2001) Purification and characterization of an aspartic protease from potato leaves. *Physiol Plant* 112:321–326.

- Guevara MG, Oliva CR, Huarte M, Daleo GR (2002) An aspartic protease with antimicrobial activity is induced after infection and wounding in intercellular fluids of potato tubers. *Eur J Plant Pathol* 108:131–137. doi: 10.1023/A:1015049629736
- Guevara MG, Oliva CR, Machinandiarena M, Daleo GR (1999) Purification and properties of an aspartic protease from potato tuber that is inhibited by a basic chitinase. *Physiol Plant* 106:164–169. doi: 10.1034/j.1399-3054.1999.106203.x
- Guiama VD, Libouga DG, Ngah E, Beka RG, Ndi KC, Maloga B, Bindzi JM, Donn P, Mbofung CM (2010) Milk-clotting potential of fruit extracts from *Solanum esculentum*, *Solanum macrocarpon* L. and *Solanum melongena*. *African J Biotechnol* 9:1797–1802. doi: 10.5897/AJB10.1079
- Guilloteau M, Laloi M, Michaux S, Bucheli P, McCarthy J (2005) Identification and characterisation of the major aspartic proteinase activity in *Theobroma cacao* seeds. *J Sci Food Agric* 85:549–562. doi: 10.1002/jsfa.1777
- Guinee TP, O'Brien B (2010) The quality of milk for cheese manufacture. In: *Technology of Cheesemaking*. Wiley-Blackwell, Oxford, UK, pp 1–67.
- Guinee TP, Wilkinson MG (1992) Rennet coagulation and coagulants in cheese manufacture. *Int J Dairy Technol* 45:94–104. doi: 10.1111/j.1471-0307.1992.tb01791.x
- Guo R, Xu X, Carole B, Li X, Gao M, Zheng Y, Wang X (2013) Genome-wide identification, evolutionary and expression analysis of the aspartic protease gene superfamily in grape. *BMC Genomics* 14:1. doi: 10.1186/1471-2164-14-554
- Gupta P, Rangan L, Ramesh TV, Gupta M (2016) Comparative analysis of contextual bias around the translation initiation sites in plant genomes. *J Theor Biol* 404:303–311. doi: 10.1016/j.jtbi.2016.06.015

- Guruprasad K, Törmäkangas K, Kervinen J, Blundell TL (1994) Comparative modelling of barley-grain aspartic proteinase: A structural rationale for observed hydrolytic specificity. *FEBS Lett* 352:131–136. doi: 10.1016/0014-5793(94)00935-X
- Habibi-Najafi MB, Lee BH (1996) Bitterness in cheese: a review. *Crit Rev Food Sci Nutr* 36:397–411.
- Hall BG (2013) Building phylogenetic trees from molecular data with MEGA. *Mol Biol Evol* 30:1229–35.
- Harboe M, Broe ML, Qvist KB (2010) The production, action and application of rennet and coagulants. In: Law BA, Tamime AY (eds) *Technology of Cheesemaking*. Wiley-Blackwell, Oxford, UK, pp 98–129.
- Harboe MK (1998) *Rhizomucor miehei* aspartic proteinases having improved properties. Springer, Boston, MA, pp 293–296.
- Hartley BS (1960) Proteolytic enzymes. *Annu Rev Biochem* 29:45–72. doi: 10.1146/annurev.bi.29.070160.000401.
- Hashim MM, Dong M, Iqbal MF, Li W, Chen X (2011a) Ginger protease used as coagulant enhances the proteolysis and sensory quality of Peshawari cheese compared to calf rennet. *Dairy Sci Technol* 91:431–440. doi: 10.1007/s13594-011-0021-x
- Hashim MM, Mingsheng D, Iqbal MF, Xiaohong C (2011b) Ginger rhizome as a potential source of milk coagulating cysteine protease. *Phytochemistry* 72:458–464.
- Hawkins CA, De Alba E, Tjandra N (2005) Solution structure of human saposin C in a detergent environment. *J Mol Biol* 346:1381–1392. doi: 10.1016/j.jmb.2004.12.045
- Heimgartner U, Pietrzak M, Geertsens R, Brodelius P, da Silva Figueiredo AC, Pais MSS (1990) Purification and partial characterization of milk clotting proteases from flowers of *Cynara cardunculus*. *Phytochemistry* 29:1405–1410. doi:

10.1016/0031-9422(90)80090-4

- Heimgartner U, Pietrzak M, Geertsen R, Brodelius P, Figueireido ACS, Pais MSS (1990) Purification and partial characterization of milk clotting. *Phytochemistry* 29:1405-1410.
- Hiraiwa N, Kondo M, Nishimura M, HaraNishimura I, Hara-Nishimura I (1997) An aspartic endopeptidase is involved in the breakdown of propeptides of storage proteins in protein-storage vacuoles of plants. *Eur J Biochem* 246:133-141.
- Hirata D, Fukui S, Yamashita I (1988) Nucleotide Sequence of the Secretable Acid Protease Gene *PEP1* in the Yeast *Saccharomycopsis fibuligera*. *Agric Biol Chem* 52:2647-2649. doi: 10.1080/00021369.1988.10869101
- Horimoto Y, Dee DR, Yada RY (2009) Multifunctional aspartic peptidase prosegments. *N Biotechnol* 25:318-324. doi: 10.1016/J.NBT.2009.03.010
- Huang J, Zhao X, Cheng K, Jiang Y, Ouyang Y, Xu C, Li X, Xiao J, Zhang Q (2013) OsAP65, a rice aspartic protease, is essential for male fertility and plays a role in pollen germination and pollen tube growth. *J Exp Bot* 64:3351-3360. doi: 10.1093/jxb/ert173
- Huang XW, Chen LJ, Luo YB, Guo HY, Ren FZ (2011) Purification, characterization, and milk coagulating properties of ginger proteases. *J Dairy Sci* 94:2259-2269. doi: 10.3168/jds.2010-4024
- Huang G-J, Huang S-S, Chen HJ, Chang YS, Chang SJ, Chang H-Y, Hsieh P-C, Chand M-J, Lin Y-C, Lin Y-H (2009) Cloning and expression of aspartic proteinase cDNA from sweet potato storage roots. *Botanical Studies* 50:149-158.
- Inan M, Aryasomayajula D, Sinha J, Meagher MM (2006) Enhancement of protein secretion in *Pichia pastoris* by overexpression of protein disulfide isomerase. *Biotechnol Bioeng* 93:771-778. doi: 10.1002/bit.20762

- Jacob M, Jaros D, Rohm H (2011) Recent advances in milk clotting enzymes. *Int J Dairy Technol* 64:14–33. doi: 10.1111/j.1471-0307.2010.00633.x
- James MNG (2004) Catalytic pathway of aspartic peptidases. In: Barrett AJ, Rawlings ND, Woessner JF (eds), *Handbook of Proteolytic Enzymes*. New York: Academic.
- Jenkins N, Curling EM (1994) Glycosylation of recombinant proteins: problems and prospects. *Enzyme Microb Technol* 16:354–64.
- Jiang XP, Yin ML, Chen P, Yang Q (2012) Constitutive expression, purification and characterization of bovine prochymosin in *Pichia pastoris* GS115. *World J Microbiol Biotechnol* 28:2087–2093. doi: 10.1007/s11274-012-1012-7
- John M, Wendeler M, Heller M, Sandhoff K, Kessler H (2006) Characterization of human saposins by NMR spectroscopy. *Biochemistry* 45:5206–5216. doi: 10.1021/bi051944+
- Joshi CP, Zhou H, Huang X, Chiang VL (1997) Context sequences of translation initiation codon in plants. *Plant Mol Biol* 35:993–1001.
- Kadek A, Tretyachenko V, Mrazek H, Ivanova L, Halada P, Rey M, Schriemer DC, Man P (2014) Expression and characterization of plant aspartic protease nepenthesin-1 from *Nepenthes gracilis*. *Protein Expr Purif* 95:121–128. doi: 10.1016/j.pep.2013.12.005
- Kageyama T (2002) Pepsinogens, progastricsins, and prochymosins: structure, function, evolution, and development. *Cell Mol Life Sci* 59:288–306.
- Kappeler SR, Van Den Brink H J, Rahbek-Nielsen H, Farah Z, Puan Z, Hansen EB, Johansen E (2006) Characterization of recombinant camel chymosin reveals superior properties for the coagulation of bovine and camel milk. *Biochem Biophys Res Commun* 342:647–654. doi: 10.1016/j.bbrc.2006.02.014

- Katsaros GI, Tavantzis G, Taoukis PS (2010) Production of novel dairy products using actinidin and high pressure as enzyme activity regulator. *Innov Food Sci Emerg Technol* 11:47–51. doi: 10.1016/J.IFSET.2009.08.007
- Kazemipour N, Salehi Inchebron M, Valizadeh J, Sepehrimanesh M (2017) Clotting characteristics of milk by *Withania coagulans*: Proteomic and biochemical study. *Int J Food Prop* 20:1290–1301. doi: 10.1080/10942912.2016.1207664
- Ke N, Berkmen M (2014) Production of disulfide-bonded proteins in *Escherichia coli*. *Curr. Protoc. Mol. Biol.* 108:16.1B.1–16.1B.21
- Kervinen J, Tobin GJ, Costa J, Waugh DS, Wlodawer A, Zdanov A (1999) Crystal structure of plant aspartic proteinase prophytepsin: Inactivation and vacuolar targeting. *EMBO J* 18:3947–3955. doi: 10.1093/emboj/18.14.3947
- Kervinen J, Wlodawer A (2013) Phytepsin. In: Rawlings ND, Salvesen G (eds) *Handbook of Proteolytic Enzymes*. Elsevier, pp 118–124.
- Kheir SEO, El Owni OAO, Abdalla MOM (2011) Comparison of quality of sudanese white cheese (Gibna Bayda) manufactured with *Solanum dubium* fruit extract and rennet. *Pakistan J Nutr* 10:106–111
- Koelsch G, Mareš M, Metcalf P, Fusek M, Maresb M, Metcalf P, Fusekat M (1994) Multiple functions of pro-parts of aspartic proteinase zymogens. *FEBS Lett* 343:6–10. doi: 10.1016/0014-5793(94)80596-2
- Kolter T, Sandhoff K (2010) Lysosomal degradation of membrane lipids. *FEBS Lett* 584:1700–1712. doi: 10.1016/j.febslet.2009.10.021
- Koontz L (2014) TCA Precipitation. *Methods Enzymol* 541:3–10.
- Kumar A, Grover S, Sharma J, Batish VK (2010) Chymosin and other

- milk coagulants: sources and biotechnological interventions. Crit Rev Biotechnol 30:243–58. doi: 10.3109/07388551.2010.483459
- Kumari M, Sharma A, Jagannadham MV (2012) Religiosin B, a milk-clotting serine protease from *Ficus religiosa*. Food Chem 131:1295–1303.
- Kumari M, Sharma A, Jagannadham MV (2010) Decolorization of crude latex by activated charcoal, purification and physico-chemical characterization of religiosin, a milk-clotting serine protease from the latex of *Ficus religiosa*. J Agric Food Chem 58:8027–8034. doi: 10.1021/jf101020u
- Kuwabara T, Suzuki K (1995) Reversible changes in conformation of the 23-kDa protein of photosystem II and their relationship to the susceptibility of the protein to a proteinase from photosystem II membranes. Plant Cell Physiol 36:495–504.
- Laemmli UK (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature 227:680–5.
- Laloi M, McCarthy J, Morandi O, Gysler C, Bucheli P (2002) Molecular and biochemical characterisation of two aspartic proteinases TcAP1 and TcAP2 from *Theobroma cacao* seeds. Planta 215:754–762. doi: 10.1007/s00425-002-0818-1
- Lee JD, Kolattukudy PE (1995) Molecular cloning of the cDNA and gene for an elastinolytic aspartic proteinase from *Aspergillus fumigatus* and evidence of its secretion by the fungus during invasion of the host lung. Infect Immun 63:3796–803.
- Lemieux L, Simard RE (1992) Bitter flavour in dairy products. II. A review of bitter peptides from caseins: their formation, isolation and identification, structure masking and inhibition. Lait 72:335–385.
- Li J, Chi Z, Wang X (2010) Cloning of the SAP6 gene of *Metschnikowia reukaufii* and its heterologous expression and characterization in *Escherichia coli*. Microbiol Res 165:173–182.

doi: 10.1016/j.micres.2008.08.003

- Libouga DG, Womeni HM, Mbofung C (2004) A milk clotting agent from the bark of *Ongokea gore* tree. Trop Sci 44:101–104. doi: 10.1002/ts.145
- Liburdi K, Emiliani Spinelli S, Benucci I, Lombardelli C, Esti M (2018) A preliminary study of continuous milk coagulation using *Cynara cardunculus* flower extract and calf rennet immobilized on magnetic particles. Food Chem 239:157–164.
- Liepinsh E, Andersson M, Ruyschaert JM, Otting G (1997) Saposin fold revealed by the NMR structure of NK-lysin. Nat Struct Biol 4:793–5.
- Lin X (2006) Universal procedure for refolding recombinant proteins. United States Patent US6583268B2.
- Lin-Cereghino J, Lin-Cereghino GP (2007) Vectors and Strains for Expression. In: Cregg JM (ed) Pichia Protocols. Humana Press, pp 11–25.
- Lin X, Koelsch G, Wu S, Downs D, Dashti A, Tang J (2000) Human aspartic protease memapsin 2 cleaves the β -secretase site of β -amyloid precursor protein. Proc Natl Acad Sci U S A 97:1456.
- Lindqvist Y, Schneider G (1997) Circular permutations of natural protein sequences: structural evidence Curr Opin Struct Biol 7:422–427.
- Linnaeus C (1753) Species plantarum 1: 107.
- Liu H, Qin Y, Huang Y, Chen Y, Cong P, He Z (2014) Direct Evaluation of the Effect of gene dosage on secretion of protein from yeast *Pichia pastoris* by expressing EGFP. J Microbiol Biotechnol 24:144–151. doi: 10.4014/jmb.1308.08010
- Llorente BE, Brutti CB, Caffini NO (2004) Purification and characterization of a milk-clotting aspartic proteinase from globe artichoke (*Cynara scolymus* L.). J Agric Food Chem 52:8182–8189. doi: 10.1021/JF049006O

- Llorente BE, Obregón WD, Avilés FX, Caffini NO, Vairo-Cavalli S (2014) Use of artichoke (*Cynara scolymus*) flower extract as a substitute for bovine rennet in the manufacture of Gouda-type cheese: Characterization of aspartic proteases. *Food Chem* 159:55–63. doi: 10.1016/j.foodchem.2014.03.007
- Lo Piero AR, Puglisi I, Petrone G (2002) Characterization of “lettucine”: A serine-like protease from *Lactuca sativa* leaves, as a novel enzyme for milk clotting. *J Agric Food Chem* 50:2439–2443. doi: 10.1021/jf011269k
- Lo Piero AR, Puglisi I, Petrone G (2011) Characterization of the purified actinidin as a plant coagulant of bovine milk. *Eur Food Res Technol* 233:517–524. doi: 10.1007/s00217-011-1543-4
- Lopes A, Teixeira G, Liberato M., Pais M., Clemente A (1998) New vegetal sources for milk clotting enzymes. *J Mol Catal B Enzym* 5:63–68. doi: 10.1016/S1381-1177(98)00059-9
- Low YH, Agboola S, Zhao J, Lim MY (2006) Clotting and proteolytic properties of plant coagulants in regular and ultrafiltered bovine skim milk. *Int Dairy J* 16:335–343. doi: 10.1016/j.idairyj.2005.03.013
- Lufrano D, Faro R, Castanheira P, Parisi G, Verissimo P, Vairo-Cavalli S, Simões I, Faro C (2012) Molecular cloning and characterization of procirsin, an active aspartic protease precursor from *Cirsium vulgare* (Asteraceae). *Phytochemistry* 81:7–18. doi: 10.1016/j.phytochem.2012.05.028
- Luo F, Jiang WH, Yang YX, Li J, Jiang MF (2016) Cloning and expression of yak active chymosin in *Pichia pastoris*. *Asian-Australasian J Anim Sci* 29:1363–70. doi: 10.5713/ajas.16.0038
- Luo J, Xiao C, Zhang H, Ren F, Lei X, Yang Z, Yu Z (2018) Characterization and milk coagulating properties of *Cynanchum otophyllum* Schneid. proteases. *J Dairy Sci* 101:2842–2850. doi: 10.3168/jds.2017-13888

- Macauley-Patrick S, Fazenda ML, McNeil B, Harvey LM (2005) Heterologous protein production using the *Pichia pastoris* expression system. *Yeast* 22:249–270. doi: 10.1002/yea.1208
- Macedo AC, Malcata FX (1996) Technological optimization of the manufacture of Serra cheese. *J Food Eng* 31:433–447.
- Macedo IQ, Faro CJ, Pires EM (1993) Specificity and kinetics of the milk-clotting enzyme from cardoon (*Cynara cardunculus* L.) toward bovine k-casein. *J Agric Food Chem* 41: 1537–1540.
- Macedo IQ, Faro CJ, Pires EM (1996) Caseinolytic specificity of cardosin, an aspartic protease from the cardoon *Cynara cardunculus* L.: action on bovine α - and β -casein and comparison with chymosin. *J Agric Food Chem* 44:42–47. doi: 10.1021/jf9406929
- MacKay VL, Welch SK, Insley MY, Manney TR, Holly J, Saari GC, Parker ML (1988) The *Saccharomyces cerevisiae* BAR1 gene encodes an exported protein with homology to pepsin. *Proc Natl Acad Sci U S A* 85:55–9.
- Mahajan RT, Chaudhari GM (2014) Plant latex as vegetable source for milk clotting enzymes and their use in cheese preparation. *Int J Adv Res* 2:1173–1181.
- Mahami T, Ocloo F, Odonkor S (2012) Preliminary study on the influence of moringa seed extracts supplementation on the yield and quality of cottage cheese. *Int J Recent Trends Sci Technol* 2:4–8.
- Mahmood T, Yang PC (2012) Western blot: technique, theory, and trouble shooting. *N Am J Med Sci* 4:429–34. doi: 10.4103/1947-2714.100998
- Marchler-Bauer A, Bo Y, Han L, He J, Lanczycki CJ, Lu S, Chitsaz F, Derbyshire MK, Geer RC, Gonzales NR, Gwadz M, Hurwitz DI, Lu F, Marchler GH, Song JS, Thanki N, Wang Z, Yamashita RA, Zhang D, Zheng C, Geer LY, Bryant SH (2017)

- CDD/SPARCLE: functional classification of proteins via subfamily domain architectures. *Nucleic Acids Res* 45:D200–D203. doi: 10.1093/nar/gkw1129
- Martínez B, Miranda JM, Franco CM, Cepeda A, Rodríguez JL (2012). Development of a simple method for the quantitative determination of fatty acids in milk with special emphasis on long-chain fatty acids. *CyTA J Food* 10: 27-35.
- Marttila S, Jones BL, Mikkonen A (1995) Differential localization of two acid proteinases in germinating barley (*Hordeum vulgare*) seed. *Physiol Plant* 93:317–327. doi: 10.1111/j.1399-3054.1995.tb02235.x
- Mavi A, Terzi Z, Ozgen U, Yildirim A, Coşkun M (2004) Antioxidant properties of some medicinal plants: *Prangos ferulacea* (Apiaceae), *Sedum sempervivoides* (Crassulaceae), *Malva neglecta* (Malvaceae), *Cruciata taurica* (Rubiaceae), *Rosa pimpinellifolia* (Rosaceae), *Galium verum* subsp. *verum* (Rubiaceae), *Urtica dioica* (Urticaceae). *Biol Pharm Bull* 27(5):702-5.
- Mawa S, Husain K, Jantan I (2013) *Ficus carica* L. (Moraceae): Phytochemistry, traditional uses and biological activities. Evidence-based complementary and alternative medicine: eCAM. 2013: 974256. <http://dx.doi.org/10.1155/2013/974256>
- Mazorra-Manzano MA, Moreno-Hernández JM, Ramírez-Suarez JC, Torres-Llanez M de J, González-Córdova AF, Vallejo-Córdova B (2013a) Sour orange *Citrus aurantium* L. flowers: A new vegetable source of milk-clotting proteases. *LWT-Food Sci Technol* 54:325–330. doi: 10.1016/j.lwt.2013.07.009
- Mazorra-Manzano MA, Perea-Gutiérrez TC, Lugo-Sánchez ME, Ramírez-Suarez JC, Torres-Llanez MJ, González-Córdova AF, Vallejo-Cordoba B (2013b) Comparison of the milk-clotting properties of three plant extracts. *Food Chem* 141:1902–1907. doi: 10.1016/j.foodchem.2013.05.042

- Mazorra-Manzano MA, Yada RY (2008) Expression and characterization of the recombinant aspartic proteinase A1 from *Arabidopsis thaliana*. *Phytochemistry* 69:2439–2448. doi: 10.1016/j.phytochem.2008.07.012
- Mendieta JR, Pagano MR, Muñoz FF, Daleo GR, Guevara MG, Muñoz FF, Daleo GR, Guevara Correspondence MG, Guevara MG (2006) Antimicrobial activity of potato aspartic proteases (StAPs) involves membrane permeabilization. *Microbiology* 152:2039–2047. doi: 10.1099/mic.0.28816-0
- Mennane Z, Khedid K, Zinedine A, Lagzouli M, Ouhssine M, Elyachioui M (2007) Microbial characteristics of Klila and Jben traditionnal moroccan cheese from raw cow ' s milk. *World* 2:23–27.
- Mignone F, Gissi C, Liuni S, Pesole G (2002) Untranslated regions of mRNAs. *Genome Biol* 3:reviews0004. doi: 10.1186/gb-2002-3-3-reviews0004.
- Morales ME, Kalinna BH, Heyers O, Mann VH, Schulmeister A, Copeland CS, Loukas A, Brindley PJ (2004) Genomic organization of the *Schistosoma mansoni* aspartic protease gene, a platyhelminth orthologue of mammalian lysosomal cathepsin D. *Gene* 338:99–109. doi: 10.1016/j.gene.2004.05.017
- Munford RS, Sheppard PO, O'Hara PJ (1995) Saposin-like proteins (SAPLIP) carry out diverse functions on a common backbone structure. *J Lipid Res* 36:1653–63
- Muñoz FF, Mendieta JR, Pagano MR, Paggi RA, Daleo GR, Guevara MG (2010) The swaposin-like domain of potato aspartic protease (StAsp-PSI) exerts antimicrobial activity on plant and human pathogens. *Peptides* 31:777–785. doi: 10.1016/j.peptides.2010.02.001
- Murakami S, Kondo Y, Nakano T, Sato F (2000) Protease activity of CND41, a chloroplast nucleoid DNA-binding protein, isolated

- from cultured tobacco cells. FEBS Lett 468:15–8.
- Mustafa WA, Sulieman AME, Abdelgadir WS, Elkhaila EA (2013) Chemical composition of the white cheese produced at household level in Dueim area, White Nile state, Sudan. J Food Nutr Disord. doi: 10.4172/2324-9323.1000108
- Mutlu A, Chen X, Reddy SM, Gal S (1999) The aspartic proteinase is expressed in *Arabidopsis thaliana* seeds and localized in the protein bodies. Seed Sci Res Seed 9:75–84 doi: 10.1017/S0960258599000082
- Mutlu A, Pfeil JE, Gal S (1998) A probarley lectin processing enzyme purified from *Arabidopsis thaliana* seeds. Phytochemistry 47:1453–9.
- Nakano T, Murakami S, Shoji T, Yoshida S, Yamada Y, Sato F (1997) A novel protein with DNA binding activity from tobacco chloroplast nucleoids. Plant Cell Online 9:1673–1682. doi: 10.1105/tpc.9.9.1673
- Nasr AIAM, Mohamed Ahmed IA, Hamid OIA (2016) Characterization of partially purified milk-clotting enzyme from sunflower (*Helianthus annuus*) seeds. Food Sci Nutr 4:733–41. doi: 10.1002/fsn3.338
- Néstor GM, Dely Rubí CG, Héctor JC (2012) Exploring the milk-clotting properties of a plant coagulant from the berries of *S. elaeagnifolium* var. Cavanilles. J Food Sci 77:C89–C94. doi: 10.1111/j.1750-3841.2011.02468.x
- Nielsen M, Lundegaard C, Lund O, Petersen TN (2010) CPHmodels-3.0-remote homology modeling using structure-guided sequence profiles. Nucleic Acids Res 38:W576–W581. doi: 10.1093/nar/gkq535
- Noseda DG, Recúpero MN, Blasco M, Ortiz GE, Galvagno MA (2013) Cloning, expression and optimized production in a bioreactor of bovine chymosin B in *Pichia* (*Komagataella*)

- pastoris under *AOX1* promoter. *Protein Expr Purif* 92:235–244. doi: 10.1016/j.pep.2013.08.018
- Nouani A, Dako E, Morsli A, Belhamiche N, Belbraouet S, Bellal M., Dadie A (2009) Characterization of the purified coagulant extracts derived from artichoke flowers (*Cynara scolymus*) and from the fig tree (*Ficus carica*) in light of their use in the manufacture of traditional cheeses in Algeria. *J. Food Technol.* 7:20–29.
- Núñez M, Pozo BF Del, Rodríguez-Marin MA, Gaya P, Medina M (1991) Effect of vegetable and animal rennet on chemical, microbiological, rheological and sensory characteristics of La Serena cheese. *J Dairy Res* 58:511. doi: 10.1017/S0022029900030120
- O'Mahony JA, Sousa MJ, Mcsweeney PLH (2003) Proteolysis in miniature Cheddar-type cheeses made using blends of chymosin and *Cynara cardunculus* proteinases as coagulant. *Int J Dairy Technol* 56:52–58. doi: 10.1046/j.1471-0307.2003.00078.x
- Olsen V, Cawley NX, Brandt J, Egel-Mitani M, Loh YP (1999) Identification and characterization of *Saccharomyces cerevisiae* yapsin 3, a new member of the yapsin family of aspartic proteases encoded by the YPS3 gene. *Biochem J* 339 (Pt 2):407–11.
- Ordiales E, Martín A, Benito MJ, Fernández M, Casquete R, de Guía Córdoba M (2014) Influence of the technological properties of vegetable rennet (*Cynara cardunculus*) on the physicochemical, sensory and rheological characteristics of “Torta del Casar” cheese. *Int J Dairy Technol* 67:402–409. doi: 10.1111/1471-0307.12129
- Ordiales E, Martín A, José Benito M, Hernández A, Ruiz-Moyano S, de Guía Córdoba M (2012) Technological characterisation by free zone capillary electrophoresis (FCZE) of the vegetable rennet (*Cynara cardunculus*) used in "Torta del Casar" cheese-

- making. Food Chem 133:227–235. doi: 10.1016/j.foodchem.2012.01.012
- Ortega-Olivencia A, Alcaraz JA (2007) *Galium* L., Flora Iberica.
- Oseni OA, Ekperigin MM (2013) Partial characterization of proteolytic and milk clotting enzymes in Sodom Apple *Calotropis procera* (Ait.) R.Br. (Asclepiadaceae) Plant Am J Biochem Mol Biol 3:256–263. doi: 10.3923/ajbmb.2013.256.263
- Otani H, Matsumori M, Hosono A (1991) Purification and some properties of a milk clotting protease from the young seeds of *Albizia julibrissin*. Anim Sci Technol 62:424–432.
- Pagano MR, Mendieta JR, Muñoz FF, Daleo GR, Guevara MG (2007) Roles of glycosylation on the antifungal activity and apoplast accumulation of StAPs (*Solanum tuberosum* aspartic proteases). Int J Biol Macromol 41:512–520. doi: 10.1016/J.IJBIOMAC.2007.07.009
- Pardo MF, López LMI, Caffini NO, Natalucci CL (2001) Properties of a milk clotting protease isolated from fruits of *Bromelia balansae* Mez. Biol Chem 382:871–4. doi: 10.1515/BC.2001.107
- Park H, Kusakabe I, Sakakibara Y, Kobayashi H (2001) Autoproteolytic processing of aspartic proteinase from sunflower seeds. Biosci Biotechnol Biochem 65:702–705. doi: 10.1271/bbb.65.702
- Payie KG, Weadge JT, Tanaka T, Yada RY (2000) Purification , N-terminal sequencing and partial characterization of a novel aspartic proteinase from the leaves of *Medicago sativa* L. (alfalfa) Biotechnol Lett 22:1515–1520.
- Pereira C, Pereira S, Pissarra J (2014) Delivering of proteins to the plant Vacuole-an update. Int J Mol Sci 15:7611–7623. doi: 10.3390/ijms15057611
- Pereira C, Pereira S, Satiat-Jeunemaitre B, Pissarra J, Genomics I, Cie

- F De, Pereira C, Pereira S, Satiat-Jeunemaitre B, Pissarra J (2013) Cardosin A contains two vacuolar sorting signals using different vacuolar routes in tobacco epidermal cells. *Plant J* 76:n/a-n/a. doi: 10.1111/tpj.12274
- Pereira CI, Franco MI, Gomes AMP, Malcata FX (2011) Microbiological, rheological and sensory characterization of Portuguese model cheeses manufactured from several milk sources. *LWT-Food Sci Technol* 44:2244–2252. doi: 10.1016/J.LWT.2011.06.013
- Pereira CS, da Costa DS, Pereira S, de Moura Nogueira F, Albuquerque PM, Teixeira J, Faro C, Pissarra J (2008) Cardosins in postembryonic development of cardoon: towards an elucidation of the biological function of plant aspartic proteinases. *Protoplasma* 232:203–213. doi: 10.1007/s00709-008-0288-9
- Pezeshki A, Hesari J, Ahmadi Zonoz A, Ghambarzadeh B (2011) Influence of *Withania coagulans* protease as a vegetable rennet on proteolysis of Iranian UF White Cheese. *J Agric Sci Technol* 13:567–576.
- Pfizer Central Research (1988) International submission on chymosin. Submitted to WHO by Pfizer, Inc., Groton, CT, USA.
- Pimentel C, Van Der Straeten D, Pires E, Faro C, Rodrigues-Pousada C (2007) Characterization and expression analysis of the aspartic protease gene family of *Cynara cardunculus* L. *FEBS J* 274:2523–2539. doi: 10.1111/j.1742-4658.2007.05787.x
- Pinho O, Ferreira IMPLV, Mendes E, Oliveira BM, Ferreira M (2001) Effect of temperature on evolution of free amino acid and biogenic amine contents during storage of Azeitão cheese. *Food Chem* 75:287–291. doi: 10.1016/S0308-8146(01)00109-1
- Pino A, Prados F, Galán E, McSweeney PLH, Fernández-Salguero J (2009) Proteolysis during the ripening of goats' milk cheese

- made with plant coagulant or calf rennet. *Food Res Int* 42:324–330. doi: 10.1016/J.FOODRES.2008.12.009
- Pires E, Faro C, Macedo I, Esteves C, Morgado J, Veríssimo P, Pereira D, Gomes D (1994) Flor do cardo versus quimosina no fabrico de queijos artesanais. *Quim Aliment* 54:66–68. doi: 10.1093/aob/mct150
- Ponting CP, Russell RB (1995) Swaposins: circular permutations within genes encoding saposin homologues. *Trends Biochem Sci* 20:179–80.
- Pontual EV, Carvalho BE, Bezerra RS, Coelho LC, Napoleão TH, Paiva PM (2012) Caseinolytic and milk-clotting activities from *Moringa oleifera* flowers. *Food Chem* 135:1848–1854. doi: 10.1016/j.foodchem.2012.06.087
- Pouillet B, Huertas M, Sanchez A, Cacare P, Larriba G (1991) Microbial study of Casar de Cáceres cheese throughout ripening. *J Dairy Res* 58: 231–238.
- Poza M, Sestelo ABF, Ageitos JM, Vallejo JA, Veiga-Crespo P, Villa TG (2007) Cloning and expression of the *XPR2* gene from *Yarrowia lipolytica* in *Pichia pastoris*. *J Agric Food Chem* 55:3944–3948. doi: 10.1021/jf0633894
- Prados F, Pino A, Fernández-Salguero J (2007) Effect of a powdered vegetable coagulant from cardoon *Cynara cardunculus* in the accelerated ripening of Manchego cheese. *Int J Food Sci Technol* 42:556–561. doi: 10.1111/j.1365-2621.2006.01271.x
- Prasad BD, Creissen G, Lamb C, Chattoo BB (2009) Overexpression of rice (*Oryza sativa* L.) *OsCDR1* leads to constitutive activation of defense responses in rice and *Arabidopsis*. *Mol Plant-Microbe Interact* 22:1635–1644. doi: 10.1094/MPMI-22-12-1635
- Radlowski M (2005) Proteolytic enzymes from generative organs of flowering plants (Angiospermae). *J Appl Genet* 46:247–57.

- Radlowski M, Kalinowski A, Adamczyk J, Krolikowski Z, Bartkowiak S (1996) Proteolytic activity in the maize pollen wall. *Physiol Plantarum* 98:172-178.
- Raimbault AK, Zuily-Fodil Y, Soler A, Cruz de Carvalho MH (2013) A novel aspartic acid protease gene from pineapple fruit (*Ananas comosus*): Cloning, characterization and relation to postharvest chilling stress resistance. *J Plant Physiol* 170:1536–1540. doi: 10.1016/j.jplph.2013.06.007
- Ramalho-Santos M, Pissarra J, Veríssimo P, Pereira S, Salema R, Pires E, Faro CJ (1997) Cardosin A, an abundant aspartic proteinase, accumulates in protein storage vacuoles in the stigmatic papillae of *Cynara cardunculus* L. *Planta* 203:204–212. doi: 10.1007/s004250050183
- Ramalho-Santos M, Verissimo P, Cortes L, Samyn B, Beeumen J V., Pires E, Faro C (1998) Identification and proteolytic processing of procarnosin A. *Eur J Biochem* 255:133–138. doi: 10.1046/j.1432-1327.1998.2550133.x
- Raposo S, Domingos A (2008) Purification and characterization milk-clotting aspartic proteinases from *Centaurea calcitrapa* cell suspension cultures. *Process Biochem* 43:139–144. doi: 10.1016/j.procbio.2007.11.003
- Rawlings ND (2013) Protease families, evolution and mechanism of action. In: Brix K, Stöcker W (eds) *Proteases: Structure and function*. Springer Vienna, Vienna, pp 1–36.
- Rawlings ND, Barrett AJ (1993) Evolutionary families of peptidases. *Biochem J* 290 (Pt 1):205–18.
- Rawlings ND, Barrett AJ, Finn R (2016) Twenty years of the MEROPS database of proteolytic enzymes, their substrates and inhibitors. *Nucleic Acids Res*. doi: 10.1093/nar/gkv1118
- Rawlings ND, Barrett AJ, Thomas PD, Huang X, Bateman A, Finn RD (2018) The MEROPS database of proteolytic enzymes, their

- substrates and inhibitors in 2017 and a comparison with peptidases in the PANTHER database. *Nucleic Acids Res* 46:D624–D632. doi: 10.1093/nar/gkx1134
- Reid JR, Coolbear T, Ayersb JS, Coolbear KP (1997) The action of chymosin on k-casein and its macropeptide: Effect of pH and analysis of products of secondary hydrolysis. *Int Dairy J* 7:559–569.
- Reis PJM, Malcata FX (2011) Current state of portuguese dairy products from ovine and caprine milks. *Small Rumin Res* 101:122–133. doi: 10.1016/j.smallrumres.2011.09.032
- Reis PM, Lourenço PL, Domingos A, Clemente AF, Salomé Pais M, Malcata X (2000) Applicability of extracts from *Centaurea calcitrapa* in ripening of bovine cheese. *Int Dairy J* 10:775–780. doi: 10.1016/S0958-6946(00)00105-9
- Rezanejad H, Karbalaie-Heidari HR, Rezaei S, Yousefi R (2015) Microsciadin, a new milk-clotting cysteine protease from an endemic species, *Euphorbia microsciadia*. *Biomacromol J* 1:93–103.
- Richter C, Tanaka T, Yada RY (1998) Mechanism of activation of the gastric aspartic proteinases: pepsinogen, progastricsin and prochymosin. *Biochem J* 335:481–90.
- Rinas U, Garcia-Fruitós E, Corchero JL, Vázquez E, Seras-Franzoso J, Villaverde A (2017) Bacterial Inclusion Bodies: Discovering Their Better Half. *Trends Biochem Sci* 42:726-737.
- Rodrigo I, Vera P, Conejero V (1989) Degradation of tomato pathogenesis-related proteins by an endogenous 37-kDa aspartyl endoproteinase. *Eur J Biochem* 184:663–669. doi: 10.1111/j.1432-1033.1989.tb15064.x
- Rodrigo I, Vera P, Van Loon LC, Conejero V (1991) Degradation of tobacco pathogenesis-related proteins. *Plant Physiol* 95:616–622. doi: 10.1104/pp.95.2.616

- Roepstorff P (1994) The interplay between different mass spectrometric techniques in protein chemistry. *Biochem Soc Trans* 22:533–9.
- Romanos M (1995) Advances in the use of *Pichia pastoris* for high-level gene expression. *Curr Opin Biotechnol* 6:527–533. doi: 10.1016/0958-1669(95)80087-5
- Roncada P, Piras C, Soggiu A, Turk R, Urbani A, Bonizzi L (2012) Farm animal milk proteomics. *J Proteom* 75:4259-4274. doi: 10.1016/j.jprot.2012.05.028
- Rosano GL, Ceccarelli EA (2014) Recombinant protein expression in *Escherichia coli*: advances and challenges. *Front Microbiol* 5:172. doi: 10.3389/fmicb.2014.00172
- Roseiro LB, Barbosa M, Ames JM, Wilbey RA (2003) Cheesemaking with vegetable coagulants - The use of *Cynara* L. for the production of ovine milk cheeses. *Int J Dairy Technol* 56:76–85. doi: 10.1046/j.1471-0307.2003.00080.x
- Rossmann M, Schultz-Heienbrok R, Behlke J, Rimmel N, Alings C, Sandhoff K, Saenger W, Maier T (2008) Crystal structures of human saposins C and D: Implications for lipid recognition and membrane interactions. *Structure* 16:809–817. doi: 10.1016/J.STR.2008.02.016
- Runeberg-Roos P, Kervinen J, Kovaleva V, Raikhel N V, Cal S (1994) The aspartic proteinase of barley is a vacuolar enzyme that processes probarley lectin *in vitro*. *Plant Physiol* 105:321–329.
- Runeberg-Roos P, Saarma M (1998) Phytapsin, a barley vacuolar aspartic proteinase, is highly expressed during autolysis of developing tracheary elements and sieve cells. *Plant J* 15:139–145. doi: 10.1046/j.1365-313X.1998.00187.x
- Runeberg-roos P, Tormakangas K, Ostman A (1991) Primary structure of a barley-grain aspartic proteinase. A plant aspartic

- proteinase resembling mammalian cathepsin D. Eur J Biochem 202:1021–1027. doi: 10.1111/j.1432-1033.1991.tb16465.x
- Salehi M, Aghamaali MR, Sajedi RH, Asghari SM, Jorjani E (2017) Purification and characterization of a milk-clotting aspartic protease from *Withania coagulans* fruit. Int J Biol Macromol 98:847–854. doi: 10.1016/j.ijbiomac.2017.02.034
- Salmia MA (1981) Proteinase activities in resting and germinating seeds of Scots pine, *Pinus sylvestris*. Physiol Plant 53:39-47.
- Sambrook J, Fritsch EF, Maniatis T (1989) Molecular Cloning. A laboratory manual. Cold Spring Harbor NY, Cold Spring Harbor Laboratory Press.
- Sampaio PN de S, Neto H, Poejo P, Serrazina SMT, Pais MS (2010) Overexpression and characterization of cyprosin B in transformed suspension cells of *Cynara cardunculus*. Plant Cell Tissue Organ Cult 101:311–321. doi: 10.1007/s11240-010-9690-z
- Sampaio PN, Fortes AM, Cabral JM, Pais MS, Fonseca LP (2008) Production and characterization of recombinant cyprosin B in *Saccharomyces cerevisiae* (W303-1A) strain. J Biosci Bioeng 105:305–312. doi: 10.1263/jbb.105.305
- Sampaio PN, Pais MS, Fonseca LP (2014) A novel fed-batch based strategy for enhancing cell-density and recombinant cyprosin B production in bioreactors. Bioprocess Biosyst Eng 37:2515–2527. doi: 10.1007/s00449-014-1229-y
- Sanjuán E, Millán R, Saavedra P, Carmona M, Gómez R, Fernández-Salguero J (2002) Influence of animal and vegetable rennet on the physicochemical characteristics of Los Pedroches cheese during ripening. Food Chem 78:281–289. doi: 10.1016/S0308-8146(02)00098-5
- Sarmiento AC, Lopes H, Oliveira CS, Vitorino R, Samyn B, Sergeant K, Debyser G, Van Beeumen J, Domingues P, Amado F, Pires E,

- Domingues MRM, Barros MT (2009) Multiplicity of aspartic proteinases from *Cynara cardunculus* L. *Planta* 230:429–439. doi: 10.1007/s00425-009-0948-9
- Schaaf A, Reski R, Decker EL (2004) A novel aspartic proteinase is targeted to the secretory pathway and to the vacuole in the moss *Physcomitrella patens*. *Eur J Cell Biol* 83:145–152. doi: 10.1078/0171-9335-00371
- Schaller A, Ryan CA (1996) Molecular cloning of a tomato leaf cDNA encoding an aspartic protease, a systemic wound response protein. *Plant Mol Biol* 31:1073–7
- Schmidt M, Polednik C, Roller J, Hagen R (2014) *Galium verum* aqueous extract strongly inhibits the motility of head and neck cancer cell lines and protects mucosal keratinocytes against toxic DNA damage. *Oncol Rep* 32:1296–302. doi: 10.3892/or.2014.3316.
- Schulte U, El-Shagi, Zenk MH (1984) Optimization of 19 *Rubiaceae* species in cell culture for the production of anthraquinones. *Plant Cell Reports* 3: 51–54.
- Senthilkumar S, Ramasamy D, Subramanian S (2006) Isolation and partial characterisation of milk-clotting aspartic protease from *Streblus asper*. *Food Sci Technol Int* 12:103–109. doi: 10.1177/1082013206063839
- Shah MA, Mir SA, Paray MA (2014) Plant proteases as milk-clotting enzymes in cheesemaking: A review. *Dairy Sci Technol* 94:5–16. doi: 10.1007/s13594-013-0144-3
- Sharp PM, Li W (1987) The codon Adaptation Index - a measure of directional synonymous codon usage bias, and its potential applications. *Nucleic Acids Res* 15:1281–1295.
- Sheehan JJ, Oliveira JC, Kelly AL, Mc Sweeney PLH (2007) Effect of cook temperature on primary proteolysis and predicted residual chymosin activity of a semi-hard cheese manufactured using

- thermophilic cultures. Int Dairy J 17:826–834. doi: 10.1016/J.IDAIRYJ.2006.08.012
- Sidrach L, García-Cánovas F, Tudela J, Rodríguez-López JN (2005) Purification of cynarases from artichoke (*Cynara scolymus* L.): enzymatic properties of cynarase A. Phytochemistry 66:41–49. doi: 10.1016/j.phytochem.2004.10.005
- Silva AC, Nascimento TC, Silva SA, Herculano PN, Moreira KA (2013) Potential of quixaba (*Sideroxylon obtusifolium*) latex as a milk-clotting agent. Food Sci Technol 33:494–499. doi: 10.1590/S0101-20612013005000075
- Silva SV, Malcata FX (2005) Studies pertaining to coagulant and proteolytic activities of plant proteases from *Cynara cardunculus*. Food Chem 89:19–26. doi: 10.1016/j.foodchem.2004.01.074
- Silva SV, Barros RM, Malcata FX (2002) Hydrolysis of caseins by extracts of *Cynara cardunculus* precipitated by ammonium sulfate. J Food Sci 67:1746–1751. doi: 10.1111/j.1365-2621.2002.tb08717.x
- Silva SV, Malcata FX (2000) Comparative catalytic activity of two plant proteinases upon caprine caseins in solution. Food Chem 71:207–214. doi: 10.1016/S0308-8146(00)00157-6
- Silva SV, Pihlanto A, Malcata FX (2006) Bioactive peptides in ovine and caprine cheeselike systems prepared with proteases from *Cynara cardunculus*. J Dairy Sci 89:3336–3344. doi: 10.3168/jds.S0022-0302(06)72370-0
- Simões I, Faro C (2004) Structure and function of plant aspartic proteinases. Eur J Biochem 271:2067–2075. doi: 10.1111/j.1432-1033.2004.04136.x
- Simões I, Faro R, Bur D, Faro C (2007) Characterization of recombinant CDR1, an *Arabidopsis* aspartic proteinase involved in disease resistance. J Biol Chem 282:31358–31365. doi:

10.1074/jbc.M702477200

- Simões I, Mueller EC, Otto A, Bur D, Cheung AY, Faro C, Pires E (2005) Molecular analysis of the interaction between cardosin A and phospholipase D α : Identification of RGD/KGE sequences as binding motifs for C2 domains. *FEBS J* 272:5786–5798. doi: 10.1111/j.1742-4658.2005.04967.x
- Singh A, Upadhyay V, Panda AK (2015) Solubilization and refolding of inclusion body proteins. *Methods Mol Biol* 1258:283–291
- Singh AN, Dubey VK (2011) Exploring applications of procerain B, a novel protease from *Calotropis procera*, and characterization by N-terminal sequencing as well as peptide mass fingerprinting. *Appl Biochem Biotechnol* 164:573–580. doi: 10.1007/s12010-011-9158-6
- Singh SM, Panda AK (2005) Solubilization and refolding of bacterial inclusion body proteins. *J Biosci Bioeng* 99:303–310. doi: 10.1263/jbb.99.303
- Soares PM, Sampaio PN, Pais MS et al. (2008) Pharmaceutical compositions containing the enzyme cyprosin, an aspartic peptidase from *Cynara cardunculus* and its inclusion in antitumour formulations. Pub No: WO/2009/040778.
- Soares MS, Calixto FC, Planta RJ (2000) Production by yeasts of aspartic proteinases from plant origin with sheep's, cow's, goat's milk, etc. clotting and proteolytic activity. Pub No: WO/2000/075283.
- Soni A, Rakha R (2017) Studies on the production and qualitative characteristics evaluation of cheese using pineapple [*Ananas comosus* (L) Merr.]. *Int J Curr Res* 9:60989–60998.
- Sørensen H, Mortensen K (2005) Soluble expression of recombinant proteins in the cytoplasm of *Escherichia coli*. *Microb Cell Fact* 4:1. doi: 10.1186/1475-2859-4-1

- Sousa M. J, Ardö Y, Mcsweeney PLH (2001) Advances in the study of proteolysis during cheese ripening. *Int Dairy J* 11:327–345. doi: 10.1016/S0958-6946(01)00062-0
- Sousa MJ, Malcata FX (1997) Comparison of plant and animal rennets in terms of microbiological, chemical, and proteolysis characteristics of ovine cheese. *J Agric Food Chem* 45:74–81. doi: 10.1021/jf9506601
- Sousa MJ, Malcata FX (1998) Proteolysis of ovine and caprine caseins in solution by enzymatic extracts from flowers of *Cynara cardunculus*. *Enzyme Microb Technol* 22:305–314. doi: 10.1016/S0141-0229(97)00173-7
- Sousa MJ, Malcata FX (1996) Effects of processing conditions on the caseinolytic activity of crude extracts of *Cynara cardunculus* L. *Food Sci Technol Int* 2:255–263. doi: 10.1177/108201329600200408
- Spohner SC, Müller H, Quitmann H, Czermak P (2015) Expression of enzymes for the usage in food and feed industry with *Pichia pastoris*. *J Biotechnol* 202:118–134. doi: 10.1016/J.JBIOTEC.2015.01.027
- Srivastava AK, Lu Y, Zinta G, Lang Z, Zhu J-K (2017) UTR-dependent control of gene expression in plants. *Trends Plant Sci* xx:1–12. doi: 10.1016/j.tplants.2017.11.003
- Stachowiak D, Wilimowska-Pelc A, Kołaczkowska M, Polanowski A, Wilusz T, Larsen LB (1994) Aspartic proteinase from the seeds of figleaf gourd (*Cucurbita ficifolia*). *Acta Biochim Pol* 41:181–2.
- Szczesniak AS (2002) Texture is a sensory property. *Food Qual Prefer* 13:215–225. doi: 10.1016/S0950-3293(01)00039-8
- Szecs PB (1992) The aspartic proteases. *Scand J Clin Lab Invest Suppl* 210:5–22.

- Takahashi K, Athauda S, Matsumoto K, Rajapakshe S, Kuribayashi M, Kojima M, Kubomura-Yoshida N, Iwamatsu A, Shibata C, Inoue H (2005) Nepenthesin, a unique member of a novel subfamily of aspartic proteinases: enzymatic and structural characteristics. *Curr Protein Pept Sci* 6:513–525. doi: 10.2174/138920305774933259
- Takahashi K, Niwa H, Yokota N, Kubota K, Inoue H (2008) Widespread tissue expression of nepenthesin-like aspartic protease genes in *Arabidopsis thaliana*. *Plant Physiol Biochem* 46:724–729. doi: 10.1016/j.plaphy.2008.04.007
- Talib M., Abubakar MM, Jideani IA, Hassan A (2009) Use of Jiben seeds extract to manufacture soft white cheese. *Am J Appl Sci* 6:551–554.
- Tamer IM (1993) Identification and partial purification of a novel milk clotting enzyme from *Onopordum turcicum*. *Biotechnol Lett* 15:427–432. doi: 10.1007/BF00128289
- Tamura T, Terauchi K, Kiyosaki T, Asakura T, Funaki J, Matsumoto I, Misaka T, Abe K (2007) Differential expression of wheat aspartic proteinases, WAP1 and WAP2, in germinating and maturing seeds. *J Plant Physiol* 164:470–477. doi: 10.1016/J.JPLPH.2006.02.009
- Tamura K, Stecher G, Peterson D, Filipski A, Kumar S (2013) MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. *Mol Biol Evol* 30:2725–2729.
- Tang J (2004) Pepsin A In: Barrett AJ., Rawlings ND, Woessner JF (eds) *Handbook of Proteolytic Enzymes*, 2 edn p.19-28, Elsevier, London.
- Tang J, James MN, Hsu IN, Jenkins JA, Blundell TL (1978) Structural evidence for gene duplication in the evolution of the acid proteases. *Nature* 271:618–21.
- Tavaria FK, Sousa MJ, Domingos A, Malcata FX et al. (1997)

- Degradation of caseins from milk of different species by extracts of *Centaurea calcitrapa*. J Agric Food Chem 45:3760–3765. doi: 10.1021/jf970095x
- Tavaria FK, Sousa MJ, Malcata FX (2001) Storage and lyophilization effects of extracts of *Cynara cardunculus* on the degradation of ovine and caprine caseins. Food Chem 72:79–88. doi: 10.1016/S0308-8146(00)00213-2
- Teixeira G, Santana AR, Pais MS, Clemente A (2000) Enzymes of *Opuntia ficus-indica* (L.) Miller with potential industrial applications-I. Appl Biochem Biotechnol 88:299–312. doi: 10.1385/ABAB:88:1-3:299
- Tejada L, Abellán A, Cayuela JM, Martínez-Cacha A, Fernández-Salguero J (2008) Proteolysis in goats' milk cheese made with calf rennet and plant coagulant. Int Dairy J 18:139–146. doi: 10.1016/J.IDAIRYJ.2007.08.010
- Terauchi K, Asakura T, Nishizawa NK, Ichiro A, Ae M, Abe K (2004) Characterization of the genes for two soybean aspartic proteinases and analysis of their different tissue-dependent expression. Planta 218:947-957.
- Terauchi K, Asakura T, Ueda H, Tamura T, Tamura K, Matsumoto I, Misaka T, Hara-Nishimura I, Abe K (2006) Plant-specific insertions in the soybean aspartic proteinases, soyAP1 and soyAP2, perform different functions of vacuolar targeting. J Plant Physiol 163:856–862. doi: 10.1016/j.jplph.2005.08.007
- Timotijević GS, Milisavljević MD, Radović SR, Konstantinović MM, Maksimović VR (2010) Ubiquitous aspartic proteinase as an actor in the stress response in buckwheat. J Plant Physiol 167:61–68. doi: 10.1016/j.jplph.2009.06.017
- Timotijević GS, Radović SR, Maksimović VR (2003) Characterization of an aspartic proteinase activity in buckwheat (*Fagopyrum esculentum* Moench) seeds. J Agric Food Chem

51:2100–2104. doi: 10.1021/jf026043b

- Tormakangas K, Törmäkangas K, Hadlington JL, Pimpl P, Hillmer S, Brandizzi F, Teeri TH, Denecke J (2001) A vacuolar sorting domain may also influence the way in which proteins leave the endoplasmic reticulum. *Plant Cell Online* 13:2021–2032. doi: 10.1105/tpc.13.9.2021
- Tormakangas K, Kervinen J, Östman A, Teeri T, Tormakangas K, Kervinen J, Ostman A, Teeri T (1994) Tissue-specific localization of aspartic proteinase in developing and germinating barley grains. *Planta* 195:116–125. doi: 10.1007/BF00206299
- Tripathi P, Tomar R, Jagannadham MV (2011) Purification and biochemical characterisation of a novel protease streblin. *Food Chem* 125:1005–1012.
- Twining SS (1984) Fluorescein isothiocyanate-labeled casein assay for proteolytic enzymes. *Anal Biochem* 143:30–4.
- Uchikoba T, Kaneda M (1996) Milk-clotting activity of cucumisin, a plant serine protease from melon fruit. *Appl Biochem Biotechnol* 56:325–330. doi: 10.1007/BF02786962
- Umezawa H, Aoyagi T, Morishima H, Matsuzaki M, Hamada M (1970) Pepstatin, a new pepsin inhibitor produced by *Actinomycetes*. *J Antibiot (Tokyo)* 23:259–62.
- Upadhyay VK, McSweeney PLH, Magboul AAA, Fox PF (2004) Proteolysis in cheese during ripening. In *Cheese*: Fox PF, McSweeney PLH, Cogan TM, Guinee TP (eds) Chemistry, Physics and Microbiology. Vol1, 3rd, pp 391–433. Elsevier Academic Press, London, UK.
- Vairo-Cavalli S, Silva SV, Cimino C, Xavier Malcata F, Priolo N (2008) Hydrolysis of caprine and ovine milk proteins, brought about by aspartic peptidases from *Silybum marianum* flowers. *Food Chem* 106:997–1003. doi: 10.1016/j.foodchem.2007.07.015

- Vairo-Cavalli S, Claver S, Priolo N, Natalucci C (2005) Extraction and partial characterization of a coagulant preparation from *Silybum marianum* flowers. Its action on bovine caseinate. J Dairy Res 72:271–275. doi: 10.1017/S0022029905000749
- Vairo-Cavalli S, Lufrano D, Colombo ML, Priolo N (2013) Properties and applications of phytepsins from thistle flowers. Phytochemistry 92:16–32. doi: 10.1016/j.phytochem.2013.04.013
- Vallejo JA, Ageitos JM, Poza M, Villa TG (2012) Short communication: A comparative analysis of recombinant chymosins. J Dairy Sci 95:609–613. doi: 10.3168/jds.2011-4445
- Vallejo JA, Ageitos JM, Poza M, Villa TG (2008) Cloning and expression of buffalo active chymosin in *Pichia pastoris* cloning and expression of buffalo active chymosin in *Pichia pastoris*. 56:10606–10610. doi: 10.1021/jf802339e
- Vallejo LF, Rinas U (2004) Strategies for the recovery of active proteins through refolding of bacterial inclusion body proteins. Microb Cell Fact 3:11. doi: 10.1186/1475-2859-3-11
- van den Brink HJ, Petersen SG, Rahbek-Nielsen H, Hellmuth K, Harboe M (2006) Increased production of chymosin by glycosylation. J Biotechnol 125:304–310. doi: 10.1016/j.jbiotec.2006.02.024
- van der Hoorn RAL (2008) Plant Proteases: From phenotypes to molecular mechanisms. Annu Rev Plant Biol 59:191–223. doi: 10.1146/annurev.arplant.59.032607.092835
- Verissimo P, Esteves C, Faro C, Pires E (1995) The vegetable rennet of *Cynara cardunculus* L. contains two proteinases with chymosin and pepsin-like specificities. Biotechnol Lett 17:621–626.
- Verissimo P, Faro C, Moir AJG, Lin Y, Tang J, Pires E, Verissimo P, Faro C, Moir AJG, Lin Y, Tang J, Pires E (1996) Purification,

- characterization and partial amino acid sequencing of two new aspartic proteinases from fresh flowers of *Cynara cardunculus* L. Eur J Biochem 235:762–768. doi: 10.1111/j.1432-1033.1996.00762.x
- Vieira M, Pissarra J, Veríssimo P, Castanheira P, Costa Y, Pires E, Faro C (2001) Molecular cloning and characterization of cDNA encoding cardosin B, an aspartic proteinase accumulating extracellularly in the transmitting tissue of *Cynara cardunculus* L. Plant Mol Biol 45:529–539. doi: 10.1023/A:1010675015318
- Vioque M, Gómez R, Sánchez E, Mata C, Tejada L, Fernández-Salguero J (2000) Chemical and microbiological characteristics of Ewes' milk cheese manufactured with extracts from flowers of *Cynara cardunculus* and *Cynara humilis* as coagulants. J Agric Food Chem 48:451–456. doi: 10.1021/jf990326v
- Wang L, Wang M, Li Q, Cai T, Jiang W (2008) Partial properties of an aspartic protease in bitter melon (*Momordica charantia* L.) fruit and its activation by heating. Food Chem 108:496–502. doi: 10.1016/j.foodchem.2007.10.085
- Wang N, Wang KY, Li G, Guo W, Liu D (2015) Expression and characterization of camel chymosin in *Pichia pastoris*. Protein Expr Purif 111:75–81. doi: 10.1016/j.pep.2015.03.012
- Waterham HR, Digan ME, Koutz PJ, Lair S V, Cregg JM (1997) Isolation of the *Pichia pastoris* glyceraldehyde-3-phosphate dehydrogenase gene and regulation and use of its promoter. Gene 186:37–44.
- Whelan S, Goldman N (2001) A general empirical model of protein evolution derived from multiple protein families using a maximum-likelihood approach. Mol Biol Evol 18:691–699.
- White PC, Cordeiro MC, Arnold D, Brodelius PE, John K (1999) Processing, activity, and inhibition of recombinant cyprosin, an aspartic proteinase from cardoon (*Cynara cardunculus*). J Biol

- Chem 274:16685–16693. doi: 10.1074/jbc.274.24.16685
- Wiederstein M, Sippl MJ (2007) ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. *Nucleic Acids Res* 35:W407–10. doi: 10.1093/nar/gkm290
- Wilimowska-Pelc A, Polanowski A, Kołaczowska MK, Wieczorek M, Wilusz T (1983) Aspartyl proteinase from cucumber (*Cucumis sativus*) seeds. Preparation and characteristics. *Acta Biochim Pol* 30:23–31.
- Xia Y, Suzuki H, Borevitz J, Blount J, Guo Z, Patel K, Dixon RA, Lamb C (2004) An extracellular aspartic protease functions in Arabidopsis disease resistance signaling. *EMBO J* 23:980–8. doi: 10.1038/sj.emboj.7600086
- Yadav RP, Patel AK, Jagannadham MV (2012) Neriifolin S, a dimeric serine protease from *Euphorbia neriifolia* Linn.: Purification and biochemical characterisation. *Food Chem* 132:1296–1304. doi: 10.1016/j.foodchem.2011.11.107
- Yadav RP, Patel AK, Jagannadham MV (2011) Purification and biochemical characterization of a chymotrypsin-like serine protease from *Euphorbia neriifolia* Linn. *Process Biochem* 46:1654–1662. doi: 10.1016/j.procbio.2011.05.013
- Yamaguchi H, Miyazaki M (2014) Refolding techniques for recovering biologically active recombinant proteins from inclusion bodies. *Biomolecules* 4:235–51. doi: 10.3390/biom4010235
- Yang Z, Zhang Z (2018) Engineering strategies for enhanced production of protein and bio-products in *Pichia pastoris*: A review. *Biotechnol Adv* 36:182–195. doi: 10.1016/j.biotechadv.2017.11.002
- Yanisch-Perron C, Vieira J, Messing J (1985) Improved M13 phage cloning vectors and host strains: nucleotide sequences of the

M13amp18 and pUC19 vectors. *Gene* 33: 103-119

- Yao X, Xiong W, Ye T, Wu Y (2012) Overexpression of the aspartic protease ASPG1 gene confers drought avoidance in *Arabidopsis*. *J Exp Bot* 63:2579–93. doi: 10.1093/jxb/err433
- Yegin S, Dekker P (2013) Progress in the field of aspartic proteinases in cheese manufacturing: Structures, functions, catalytic mechanism, inhibition, and engineering. *Dairy Sci Technol* 93:565–594. doi: 10.1007/s13594-013-0137-2
- Yegin S, Fernandez-Lahore M (2013) A Thermolabile aspartic proteinase from *Mucor mucedo* DSM 809: gene identification, cloning, and functional expression in *Pichia pastoris*. *Mol Biotechnol* 54:661–672. doi: 10.1007/s12033-012-9608-6
- Zhang AL, Luo JX, Zhang TY, Pan YW, Tan YH, Fu CY, Tu FZ (2009) Recent advances on the GAP promoter derived expression system of *Pichia pastoris*. *Mol Biol Rep* 36:1611–1619. doi: 10.1007/s11033-008-9359-4
- Zhang Y, Wang H, Tao L, Huang A (2015) Milk-clotting mechanism of *Dregea sinensis* Hemsl. protease. *J Dairy Sci* 98:8445–53. doi: 10.3168/jds.2015-9851
- Zhao R, Chen Z, Jia G, Li J, Cai Y and Shao X (2011) Protective effects of diosmetin extracted from *Galium verum* L. on the thymus of U14-bearing mice. *Can J Physiol Pharmacol* 89: 665-673
- Zhao CC, Shao JH, Li X, Xu J, Wang JH (2006) A new anthraquinone from *Galium verum* L. *Nat Prod Res* 20:981-4.
- Zhao C, Shao J, Cao D, Zhang Y and Li X (2009) Chemical constituents of *Galium verum*. *Zhongguo Zhong Yao Za Zhi* 34: 2761-2764.
- Zhao CC, Shao JH, Li X, Kang XD, Zhang YW, Meng DL, Li N (2008) Flavonoids from *Galium verum* L. *J Asian Nat Prod Res*

10: 613-617.

Zhu T, Guo M, Tang Z, Zhang M, Zhuang Y, Chu J, Zhang S (2009) Efficient generation of multi-copy strains for optimizing secretory expression of porcine insulin precursor in yeast *Pichia pastoris*. J Appl Microbiol 107:954–963. doi: 10.1111/j.1365-2672.2009.04279.x







PUBLICACIONES

8. PUBLICACIONES

Este trabajo ha originado las siguientes publicaciones:

- ✓ Feijoo-Siota L, Villa TG (2011) Native and biotechnologically engineered plant proteases with industrial applications. Food Bioprocess Tech 4: 1066–1088.
<https://link.springer.com/article/10.1007/s11947-010-0431-4>
- ✓ Feijoo-Siota L, Blasco L, Rodríguez-Rama JL, Barros-Velázquez J, Miguel Td, Sánchez-Pérez A, Villa TG (2014) Recent patents on microbial proteases for the dairy industry. Recent Adv DNA Gene Seq 8:44-55.
https://www.researchgate.net/publication/270662199_Recent_Patents_on_Microbial_Proteases_for_the_Dairy_Industry
- ✓ Feijoo-Siota L, Rama JLR, Sánchez-Pérez A, Villa TG (2018) Expression, activation and processing of a novel plant milk-clotting aspartic protease in *Pichia pastoris*. J Biotechnol 268:28-39. doi: 10.1016/j.jbiotec.2018.01.006
<https://www.sciencedirect.com/science/article/pii/S0168165618300075?via%3Dihub>



FE DE ERRATAS

- 1) En la **página 89** (Figura 16), en la imagen, donde dice pMgv1r debe decir pMgv1f y viceversa.
- 2) En la **página 133** (Figura 23, B/), en el interior del exón E_A del gen gv1AP y del exón E₇ del gen oryzasin, debe figurar el número 66 en lugar del número 65.
- 3) En la **página 139** (Figura 27), en la parte inferior de la imagen, en el interior del rectángulo verde debe figurar Sap_B1 en lugar de Sap_B2 y en el interior del rectángulo granate debe figurar Sap_B2 en lugar de Sap_B1.

